ADVANCES IN ZERO-BASED CONSISTENT DECONVOLUTION AND EVALUATION OF HUMAN SENSORY-MOTOR FUNCTION.

A thesis presented for the degree of
Doctor of Philosophy in Electrical and Electronic Engineering
at the University of Canterbury, Christchurch, New Zealand.

March 1994

Russell William Watson
B.E. (Hons I)

ABSTRACT

Work in two separate research areas is presented. The two areas are the development of a new two-dimensional (2-D) zero-based deconvolution algorithm and the evaluation of human sensory-motor (S-M) function in a 2-D environment.

In Part One of this thesis the new algorithm which solves the 2-D standard deconvolution problem (i.e. deconvolution of a single contaminated blurred image when the point spread function (psf) is known *a priori*) is described. The algorithm performs the deconvolution by separating a subset of the *zeros* of the Z transform of the convolution into two further subsets. This separation is performed by using the zeros of the psf to recognise and eliminate, from the set of convolution zeros, the set of zeros that pertain to the psf actually responsible for the blurring. The remaining subset of zeros characterise the original image and an estimate of the original image can be reconstructed from these zeros by using an algebraic technique. As this technique does not explicitly use the psf to generate the image estimate (unlike existing deconvolution techniques such as inverse and Wiener filtering) it was anticipated that the quality of the reconstructions obtained would not be strongly dependent on the psf estimate accuracy.

The performance of the new zero-based deconvolution algorithm has been extensively assessed. It has been found that the algorithm can successfully deblur small (original image sizes $< 16\times16$ pixels), high SNR ($\geq$ 40 dB) blurred images. On comparing the quality of the reconstructions provided by this new algorithm to those provided by the Wiener filter, it was found that the image estimate quality of the new algorithm was less dependent on the psf accuracy than that of the Wiener filter. The effects of contamination on the position of the zeros in Z space has also been investigated. In the situation where the contamination is band-limited, it is shown that specific zeros are likely to be affected more by the contamination than others. It is shown that this knowledge can be used, in the band-limited case, to improve the quality of the reconstructions provided by the new algorithm. In comparison to existing techniques, such as the Wiener filter, this algorithm is computationally very expensive and this computational expense becomes prohibitive as the image size increases. It appears that the section of the algorithm responsible for the large computational expense and which limits the ability of the algorithm to obtain useful reconstructions from larger images is the reconstruction section, as opposed to the zero recognition and separation section. Possible techniques that could be used to improve the algorithm's performance in these two areas are described.

The phase retrieval problem is a problem very similar to deconvolution. The success of the new algorithm at solving the deconvolution problem prompted the development of a zero-based algorithm which can be used to assist the convergence of one of the iterative algorithms typically used to solve the phase retrieval problem. It is shown that the zero-based algorithm can successfully assist the iterative algorithm in obtaining a solution, however the zero-based algorithm's enormous computational cost results in no overall computational saving.

In Part Two two new 2-D eye-arm pursuit tracking tasks, which have been developed to facilitate the investigation of human S-M function in a 2-D environment, are described. The tracking tasks assess a subject's ability to accurately follow a moving visual target with a pointer controlled by the coordinated movement of an upper limb. One of the tracking tasks uses a smoothly changing pseudo-random target signal and the other uses a signal consisting of temporally and spatially unpredictable steps. The 2-D tasks were developed with the anticipation that they might provide more sensitive measures than their 1-D counterparts of, as well as additional information on, both abnormal and normal S-M function. Two trials where conducted to determine what benefits the new 2-D tracking tasks could provide. It was found that the 2-D tasks could successfully demonstrate the deficits known to occur in the Parkinsonian S-M system and that the 2-D tasks could detect these deficits with a marginally greater sensitivity than 1-D tasks. It was also found that on going from a 1-D to a 2-D tracking task, the observed degradation in tracking performance consists of two components. The first component is due to an increase in the S-M system processing delay when tracking multidimensional signals and the second is due to a reduction in tracking accuracy with any increase in average task displacement and/or velocity. An age dependent difference between the subject's performance in the vertical and horizontal directions of the 2-D tracking task was also observed.

One of the main performance measures used with tracking tasks is the time delay between the target moving and the S-M system responding. An evaluation of the accuracy of the various techniques which can be used to calculate the time delay has been conducted. Of the five techniques tested, it was found that the standard cross-correlation technique or the least squares time delay estimation technique could provide the most accurate estimates of the time delay. The reasons for the inaccuracies in the time delay estimates provided by some of the other techniques in the comparison are also explained.

ACKNOWLEDGEMENTS

First and foremost I thank my wife and good friend Julie for her love, patience and support when I needed it most. This thesis is testimony to the sacrifices that Julie has also made and the hard work she has put in.

My sincere thanks also go to my supervisor Dr Philip Bones and my associate supervisor Dr Richard Jones, who have provided me with encouragement and guidance during the course of my studies. Phil and Richard have an almost contagious enthusiasm for research which has rubbed off on me and, I believe, added to the quality of the work presented here. They both are the type of supervisor who takes a real interest in their student's work and puts in a significant amount of hard work to ensure their students are successful - for this I thank them.

I am grateful to my original supervisor, the late Professor Richard H T Bates, for his encouragement to undertake PhD studies and for the initial guidance he provided.

My colleagues Charles Parker and Brenda Satherley deserve special mention. I thank them for providing me with an enormous amount of encouragement, acting as a sounding board for some of my ideas, and for all the help they have given me.

The past and present members of R6 and R7 I thank you for ensuring I maintained at least a degree of sanity and for the comradeship you have provided.

My gratitude also goes to all those staff in the Electrical and Electronic Engineering Department and at the Medical Physics and Bioengineering Department at Christchurch Hospital who have helped me in any way.

I gratefully acknowledge the financial support provided by the Telecom Corporation of New Zealand Ltd and the University Grants Committee.

Finally I thank my family and Julie's family for their encouragement and for putting up with my tunnel vision over the last few years.

Dedicated to the memory
of my mother

Ellen Emily Watson

1939-1973

CONTENTS

PREFACE

Human beings have an unassailable desire to gain a complete understanding of themselves and the world around them. To obtain the information required to develop this understanding, people use many different methodologies to observe and probe objects of interest. The information is generally in the form of measurements of some physical quantity, which indirectly or directly describes the properties of the object of interest to the observer. The original research contained within this thesis is concerned with the development of two completely different techniques which can be used to obtain information about an object. The first of these is a technique that has been developed to recover information from a degraded observation of (in principle) any object. Specifically, the first technique is concerned with recovering an estimate of an image from a blurred and contaminated version of the image when the form of the blurring is known *a priori*. In the field of imaging science there are many imaging techniques that could produce such a degraded image and so the applications for the above technique could be many and varied. The second technique is a technique which has been used to obtain information on the function of the human sensory-motor system and its major component, the brain. This technique uses what is known as a tracking task to stimulate the sensory-motor system and to observe the response. The manner in which the sensory-motor system responds to the stimulus can provide valuable information on both the normal and abnormal sensory-motor system. While these two techniques are concerned with extracting information about an object, the methodologies used in each case are vastly different and any attempt to link the two techniques, beyond the global perspective described above, would be largely artificial. For this reason, this thesis is divided into two parts, with Part One describing the research in the imaging area and Part Two describing the research in the sensory-motor area. The aim of the following paragraphs is to provide the reader with a description of how I became involved with the above research, the basic motivation for the research, an outline of the contents of this thesis, and an overview of what is original.

My first introduction to multidisciplinary research and the research of the information processing group lead by the late Professor Richard Bates, came about in 1989 during the final year of my Bachelor of Electrical Engineering Degree. Professor Bates, in conjunction with Dr Richard Jones from Christchurch Hospital, supervised my final year project, which involved developing a series of two-dimensional (2-D) tracking tasks for the assessment of sensory-motor (S-M) function. These tracking tasks were the predecessors of the set of 2-D tracking tasks described in Part Two. In 1990 I began my PhD studies and joined Professor Bates' information processing group, well known for its multidisciplinary research in the information processing area. At the time, members of the group were investigating aspects of image processing, speech processing, radio antenna engineering, biomedical engineering, and chaos theory. Professor Bates introduced me to a member of the group, Charles Parker, who through a series of weekly meetings introduced me to his field of research, which involved

using the properties of the zeros of the Fourier transform to solve the phase retrieval problem. We decided that the application of similar principles to the 2-D contaminated deconvolution problem would be a promising area of research, and so this became one of my PhD topics. Following Professor Bates' tradition of multidisciplinary research and of exposing research students to a number of research areas, my work on the use of 2-D tracking tasks to assess S-M function became another one of my initial PhD topics, along with research into aspects of chaos in oscillator circuits.

The research into aspects of chaotic oscillators progressed to a stage where I was able to simulate the output of a number of oscillators on a computer, by using numerical techniques (Runge Kutta) to solve the differential equations describing the oscillators' behaviour. Unfortunately, Professor Bates died in November 1990 and this area of research had to be curtailed as it was no longer feasible to continue. Dr Philip Bones agreed to take over my supervision and I continued my research into the two remaining areas and also began looking at modelling aspects of the Magnetoencephalogram (M.E.G.) with Dr Bones. By the middle of 1991 my research in the deconvolution and tracking task areas was providing fruitful results, and so these two areas of research became my official PhD topics and the research in the M.E.G area was terminated. My brief introduction to the M.E.G. problem was not wasted, however, as it has left me with an understanding of many aspects of electromagnetic theory.

The work in Part One of this thesis, which describes an original algorithm to deconvolve blurred and contaminated images, was motivated by the previous work of Professor Bates' research group (Lane *et al.*, 1987; Lane and Bates, 1987; Bates *et al.*, 1990). Lane and Bates had shown that, in the uncontaminated situation, it was possible to use the properties of the zeros of the Z transform of a convolution to successfully solve the phase retrieval and deconvolution problem. The principle used was that the zeros of the Z transform of the convolution can be divided into two distinct sets, one which characterises the original image and another which characterises the blurring function or the conjugate of the image. Bates *et al.* (1990) further demonstrated that for 1-D images it was possible to use the above properties to solve the contaminated deconvolution problem when an estimate of the blurring function was known *a priori*. Professor Bates, my colleagues and I saw a need to develop a new algorithm which used the zero-based techniques to solve the contaminated 2-D deconvolution problem, when the blurring function was known *a priori*. We anticipated that the new deconvolution algorithm may have the beneficial property that it may be less sensitive than existing techniques, such as the Wiener filter, to inaccuracies in the estimate of the blurring function because the new algorithm uses the blurring function estimate to recognise and remove the actual blurring. The success of this new algorithm, in solving the deconvolution problem, suggested that the algorithm may be able to assist the iterative phase retrieval algorithms in obtaining a solution. I therefore have also applied the algorithm, with success, to the solution of the phase retrieval problem.

The research into the use of 2-D tracking tasks to investigate S-M function, contained in Part Two, was motivated by the the upgrading of the existing S-M test battery at Christchurch Hospital. The 2-D tracking tasks I developed during the final year of my Bachelors degree were to be included within this upgraded battery of tests. My first task was to rewrite the 2-D tracking task software so that the tasks could be incorporated within the new battery. At this stage it was not known whether the 2-D tasks could provide users of the test battery with information, on either normal or abnormal S-M function, that was unavailable from the 1-D tasks. My research programme was to therefore investigate whether 2-D tracking tasks could demonstrate the deficits already known to occur in people with certain neurological disorders (which can be demonstrated using 1-D tasks), to determine whether there was any benefit to

be gained from using 2-D tracking tasks in preference to 1-D tasks (i.e. can 2-D tasks provide new or more sensitive information on S-M function than 1-D tasks), and to determine the differences in the mechanisms used by the S-M system to perform multidimensional tracking tasks, as opposed to single dimension tracking tasks. The above areas of research were considered to be of importance as, although numerous researchers have used 2-D tracking tasks, no-one has previously attempted to answer the above questions. The inaccuracy of the technique used, within the existing battery, to calculate the time delay between the stimulus and the S-M response prompted an additional area of research. The reason for the inaccuracy of the current technique was to be determined and alternative techniques were to be assessed for their accuracy.

This thesis is divided into eleven chapters, with Chapters 1 to 7 comprising Part One and Chapters 8 to 11 comprising Part Two. Each chapter begins with a brief introduction and an outline of its contents. The individual sections of this thesis have been numbered in accordance with the decimal numbering system and so, for example, the third subsection of the fourth section of Chapter 2 is section 2.4.3. The symbol § is used in place of the word 'section' whenever it is necessary to refer to a particular section from within the text. The tables, figures, and equations are numbered consecutively within each chapter in a fashion similar to the decimal numbering of the sections. When referring to these items from within the text, the appropriate number is preceded by the term Table, Fig., or Eqn. to signify what form the item referred to takes. A chapter-by-chapter outline of the contents of this thesis is now given.

Chapter 1 contains a brief introduction to the field of imaging science. The intention of this chapter is to convey to the reader the enormous scope of the field of imaging science, its importance, and the problems usually encountered when imaging objects.

In Chapter 2, the notation and mathematical relationships which are required in this thesis are introduced. Also described are the models used to describe the optical image formation process, the far field of an aperture antenna, and the degradation processes that corrupt images. The detailed description of the models should give the interested reader an appreciation of the assumptions that are made to obtain the expressions which describe the deconvolution and phase retrieval problems.

Chapter 3 contains a review of the existing techniques which can be used to solve the deconvolution and phase retrieval problems. The review of deconvolution techniques is limited to those which require prior knowledge of the blurring function and the review of phase retrieval techniques is limited to those iterative techniques based on the spatial/Fourier domain iterative loop. Also reviewed are some techniques which can be used to obtain estimates of the blurring function.

The concept of the zero sheet, which is the surface formed by either the zeros of the Fourier or Z transform of an image, is described in Chapter 4. The properties of zero sheets in general, as well as those of convolutions, of power spectra, and in the presence of contamination are examined. Also described are the techniques that can be used to reconstruct an image from its zero sheet. In particular the QR decomposition algorithm which is used to achieve algebraic reconstruction is described in some detail. I combine the graphical explanation of Steinhardt (1988) with the more traditional mathematical description of the technique in an attempt to simplify the description.

The new deconvolution algorithm is described in Chapter 5. Results which indicate how well the algorithm performs in the presence of convolution contamination, blurring function estimate inaccuracy, and in comparison to existing deconvolution techniques are presented.

The strengths and weaknesses of the new algorithm are extensively examined.

In Chapter 6, the application of the new algorithm to the phase retrieval problem is described. The results presented in this chapter demonstrate that the algorithm can assist the normal iterative algorithms in obtaining a solution to the phase retrieval problem, although it does so at a large computational expense. The performance of my algorithm is also compared to that recently reported by Parker and Bones (1992).

In Chapter 7, conclusions derived from the original work contained within Part One are presented and possible avenues for future research are suggested.

Chapter 8 introduces the area of human S-M research and reviews those techniques that can be used to assess the function or dysfunction of the human S-M system. The emphasis is on conveying why tracking tasks are considered to be one of the most valuable and accurate means of measuring S-M function. Also described is Christchurch Hospital's upgraded S-M test battery.

The two 2-D tracking tasks I personally developed are described first in Chapter 9. Following this description the various techniques that can be used to measure the time delay between the tracking task stimulus and the S-M system response are reviewed. The strengths and weaknesses of these techniques are analysed and the results of an experimental comparison of the accuracy of the techniques is presented.

In Chapter 10, two clinical studies which were used to assess the performance of the new 2-D tracking tasks, to compare the information provided by 2-D and 1-D tracking tasks and to determine how the sensory-motor system allocates resources when confronted by a 2-D tracking task, are described. The results from these two studies are also presented and discussed.

Chapter 11 contains the conclusions and suggestions for further research that relate to Part Two.

The original research presented in this thesis comprises the majority of Chapters 5, 6, 8, 9 and part of Chapter 4. I personally developed the majority of the software required to implement my new deconvolution and phase retrieval algorithms. Where algorithms written by others have been included within my software, this is indicated in the text. The image processing package *improc*, developed by Dr Richard Lane, has been used to generate images, to construct deconvolution filters, and to construct iterative phase retrieval algorithms. The software to implement the zero-based deconvolution and phase retrieval algorithms of Parker and Bones (1992) and Parker (1992) was provided by Charles Parker. The core software for Christchurch Hospital's S-M test battery, and the software required to implement the tests that were part of the original battery, was written by Neil Sharman and Richard Jones.

During the course of the work presented in this thesis the following papers and presentations have been prepared:

BONES, P.J., WATSON, R.W., JIANG, H. and SATHERLEY, B.L. (1992), 'Restoring images - after the Bates' tradition', In *Proc. IPENZ Ann. Conf.*, pp. 165–173.

BONES, P.J., SATHERLEY, B.L. and WATSON, R.W. (1992), 'Direct deconvolution of noisy blurred images', In *IEE 4th Conference on Image Processing and its Applications*, pp. 353–356.

BONES, P.J., SATHERLEY, B.L., PARKER, C. and WATSON, R.W. (1993), 'Non-iterative methods for image deconvolution', In IDELL, P.S. (Ed.), *Digital image recovery and synthesis II, SPIE, vol 2029*,pp. 2–13.

DALRYMPLE-ALFORD, J.C., KALDERS, A.S., JONES, R.D. and WATSON, R.W. (1992), ' Parkinson's disease and working Memory', *Int. J. Neurosci. (Abstract)*, Vol. 71, p. 114.

DALRYMPLE-ALFORD, J.C., KALDERS, A.S., JONES, R.D. and WATSON, R.W. (1994), 'A central executive deficit in patients with Parkinson's disease', *J. Neurol. Neurosurg. Psychiatry, In Press.*

JONES, R., SHARMAN, N., WATSON, R. and MUIR, S. (1992), 'A PC-based battery of tests for measurement of sensory-motor function in brain disorders', In *Proc. Ann. Conf. of N.Z. Branch of Australasian College of Phys. Sci. Eng. Med., (Abstract)*, Christchurch.

JONES, R., SHARMAN, N., WATSON, R. and MUIR, S. (1993), 'A PC-based battery of tests for quantitative assessment of upper-limb sensory-motor function in brain disorders', In *Proc. 15th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, San Diego, Vol. 15, pp. 1414–1415.

JONES, R.D. and WATSON, R.W. (1993), 'A two-dimensional step tracking task and its application to Parkinson's disease', In *Proc. 15th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, San Diego, Vol. 15, pp. 1416–1417.

SCRASE, A.N., BONES, P.J., SATHERLEY, B.L. and WATSON, R.W. (1992), 'Visualisation of the image zero sheet', In *Proceedings of the 7th New Zealand Image Processing Workshop*, pp. 107–112.

WATSON, R.W. and JONES, R.D. (1990), 'The development of PC-based two-dimensional tracking tasks for the study of neurological disorders', In *Proc. Ann. Conf. of N.Z. Branch of Australasian College of Phys. Sci. Eng. Med., (Abstract).*

WATSON, R.W. and BONES, P.J. (1991), 'An alternative to Wiener filtering', *Proceedings of the 6th N.Z. Image Processing Workshop*, pp. 139–144.

WATSON, R.W. and BONES, P.J. (1992), 'An alternative to Wiener filtering II', In BONES, P. (Ed.), *Proceedings of the 7th N.Z. Image Processing Workshop*, pp. 71–76.

WATSON, R.W. and JONES, R.D. (1992), 'An evaluation of 2-dimensional tracking performance of subjects with Parkinson's disease', In *Proc. Ann. Conf. of N.Z. Branch of Australasian College of Phys. Sci. Eng. Med., (Abstract).*

WATSON, R.W., PARKER, C.R. and BONES, P.J. (1992), 'Demonstration of 2-D consistent deconvolution using zeros', *Optics Comm*, Vol. 93, pp. 359–365.

WATSON, R.W., JONES, R.D. and SHARMAN, N.B. , 'Two-dimensional tracking tasks for quantification of sensory-motor dysfunction and their application to Parkinson's disease', *Submitted to IEEE Trans Rehab Eng.*

Part I

Zero-based Consistent Deconvolution

CHAPTER 1

INTRODUCTION

The science of measuring emanations (e.g. electromagnetic waves, acoustic waves, elastic waves, particles, etc.) being emitted from, passing through or being reflected from an object, in order to obtain a visual representation of the physical or functional attributes of the object of interest, is known as imaging. The measured representation of the object is termed an image and is usually, although not necessarily so, a function of two spatial variables, i.e. the image is most often a two-dimensional (2-D) function. It is humankind's quest for knowledge and desire to understand themselves and their environment that has prompted the increased use and advances in imaging technology over the last century. One of the major factors which has contributed to the value of imaging technology, in both research and in life in general, is that the various imaging techniques can provide physical and functional information on objects that are obscured by other objects or are too small or too far away to be physically examined. As examples consider organs within the living human body, the structure of crystals, and stars that comprise distant galaxies; it would be either impractical or undesirable to physically examine these objects for the above respective reasons. Each of these objects and most objects of interest, however, either emit radiation, can be made to emit radiation, or behave in a predetermined manner when probed with radiation and it is therefore usually possible to obtain an image of the object by using one of the many imaging techniques. The following discussion contains an introduction to imaging science and some of the imaging applications and techniques which are currently in regular use or the subject of ongoing research. The techniques and references cited in this discussion are by no means an exhaustive list but were compiled, in conjunction with Tables 1.1 and 1.2, to give the reader an appreciation of the scope and importance of imaging science. Following this introduction the value of being able to model or describe mathematically the various image formation and degradation processes is discussed. Finally, the new image deblurring technique, which is the subject of Part One of this thesis, is introduced and the motivation for its development explained.

Imaging techniques where the emanations being emitted from an object of interest are passively observed are classified as *remote sensing* techniques. With these techniques the observer generally has no control over the form of the emanations, and hence information, being emitted from the object. In a number of cases, however, the observer does have the flexibility to observe only those emanations that provide information on the characteristic of interest. In remote sensing the principal emanations that are measured to image an object are reflected radiation, thermal radiation and radiation from radioactive decay. The most widely measured of all the above emanations is electromagnetic radiation which originates from a source outside the control of the observer (e.g., the sun) and is reflected by the object. The fields of optical astronomy (Roy and Clarke, 1982; Hyde, 1992), microscopy (Rochow and Rochow, 1980), satellite imaging (Sabins, 1978; Scorer, 1986; Le Vine *et al.*, 1989) and

archival imaging (e.g., domestic photography and videography) all rely on imaging techniques which measure reflected radiation to obtain physical information on the object of interest (see Table 1.1). The reflected radiation is usually limited to the visible, infrared, and microwave wavelengths since the majority of radiation at other wavelengths is either absorbed or reflected by the earth's atmosphere (Sabins, 1978).

IMAGING APPLICATIONS AND TECHNIQUES

Remote Sensing			
Area of Application	Imaging Technique	Type of Emanation	Information
Medicine.	Nuclear imaging using: Gamma cameras, SPECT, PET.	EM. γ rays (< 0.06 Å).	Functional & Physical.
	EEG and ECG topography (Duffy *et al.*, 1989).	EM. (< 100 Hz).	Functional.
	Microwave thermography.	EM. (1 GHZ – 10 GHz).	Physical & Functional.
Astronomy.	Optical telescopes.	EM. near ultraviolet to near infrared (0.35 μm – 1 μm).	Physical.
	Radio telescopes.	EM. radio and μwaves (15 MHz – 300 GHz).	Physical.
Microscopy.	Optical microscopes.	EM. visible (0.4 μm – 0.72 μm).	Physical.
Satellite Imaging for: Geoscience, natural hazard detection, environmental imaging, land-use classification, meteorology.	Optical cameras. Infrared cameras and detectors. Microwave radiometers.	EM. visible. EM. photographic infrared (0.7 μm - 3.0 μm), thermal infrared (3 μm – 5 μm and 8 μm – 14μm). E.M. μwaves (1 GHz – 180 GHz).	Physical.
Archival Imaging.	Photography, Cinematography, Videography.	EM. Visible.	Physical.

Table 1.1 Key: EM. = Electromagnetic. Physical information is information that describes the physical structure of the object. Functional information is information that describes how the various components within an object function, e.g. in nuclear medicine the function of organs is checked for any abnormality by imaging how a organ of interest concentrates a radiopharmaceutical.

Thermal radiation is emitted by all objects at a temperature above 0 K in the form of multispectral electromagnetic energy (Sabins, 1978). The intensity of the radiation at the various wavelengths is dependent on the temperature of the body. Thus, by measuring the thermal radiation emanating from a body it is possible to obtain an image of the temperature distribution of that body in addition to information on its physical structure. In optical and radio astronomy thermal radiation at visible and microwave wavelengths respectively is used to obtain information on the location and physical structure of extraterrestrial objects (see Table 1.1) (Kraus, 1966; Smith, 1966; Roy and Clarke, 1982; Rohlfs, 1986; Hyde, 1992). Thermal radiation at microwave wavelengths has also been used experimentally to obtain images of the temperature distribution within the human body (Barrett and Myers, 1975).

Satellite imaging systems measure thermal radiation at thermal infrared wavelengths in order to obtain images of the earth for the purposes of detecting thermal pollution, geothermal energy, volcanic activity, resource bearing materials, and for classifying the various constituents of the landscape (Sabins, 1978).

Radiation originating from radioactive decay is used extensively in nuclear medicine to noninvasively obtain information on the physical structure and functional integrity of the human body (Cameron and Skofronick, 1978; Todd Pokropek, 1980; Knoll, 1983; Jaszczak, 1988). To obtain these images, substances known to distribute themselves in the body in a predetermined manner are labelled with a radioisotope and injected into the body. The distribution of this radiopharmaceutical in the body is then imaged using one of the techniques in Table 1.1 that measure the density of the γ rays that are being emitted from the body by the decay of the isotopes. The density of the emissions and how they change with time allows both functional and physical information on the organs to be deduced. Even though in nuclear medicine the observer provides the source of the emanations this form of imaging still comes under the category of remote sensing since the emanations at the time of imaging originate from within the body and are not directed at it from some outside source.

Those imaging techniques were the interaction of emanations actively directed at the object are observed are known as *remote probing* techniques (Bates *et al.*, 1991). With remote probing techniques the observer is in the advantageous position of being able to select the emanation that is to be used to probe the object. By being able to select an emanation of an appropriate intensity, form, and wavelength, the observer has more flexibility to dictate the type and optimize the quality of the information received. The four factors that primarily influence what form the emanations take are the transmissivity of the transmission medium with respect to the emanation, the size of the object's features which must be resolved, the effect the emanations have on the object, and the type of information required. The first of the above factors, which probably dictates to the largest degree what form the emanations take, arises because the attenuation of an emanation in a medium is dependent on its form and wavelength. Hence, to obtain an image, an emanation, of an appropriate wavelength, which can traverse the transmission medium with an attenuation that still allows the interaction of the emanations with the object to be accurately measured, must be used. It is for this reason that sound-waves are used to image objects when the transmission medium is predominately liquid (e.g. SONAR (Urick, 1967)), elastic waves are used when the medium is solid (e.g. Seismic imaging (Gustavsson *et al.*, 1986)), and electromagnetic waves are used when the transmission medium is the atmosphere (e.g. RADAR (Sabins, 1978; Skolnik, 1980; Scorer, 1986))(also see Table 1.2).

The size of the features on or within an object to be imaged also has a direct bearing on the form of the emanation or the wavelength. Generally the smaller the features to be resolved the smaller the wavelength or particles comprising the emanation must be. Table 1.2 contains two prime examples of techniques where this principle has been used to achieve the required resolution, these are techniques in the fields of crystallography (Millane, 1990) and microscopy (Rochow and Rochow, 1980) where X-rays and electrons are used respectively to image very small structures. The relative size of the features in relation to the particle size or wavelength also influences whether the observer is interested in how the emanations are attenuated or scattered by the medium. When an object's features are large in comparison to the wavelength used, such as in X-ray absorption imaging techniques, the required information can be obtained directly from measurements of the attenuation of the emanations or by inverting an expression which describes how the object attenuates the emanations (e.g. standard X-ray (Cameron and Skofronick, 1978) and X-ray computerised tomography (CT)(Bates *et al.*, 1983; Robb, 1985)).

IMAGING APPLICATIONS AND TECHNIQUES

Remote Probing			
Area of Application	Imaging Technique	Type of Emanation	Information
Medicine.	X-ray absorption.	EM. X-rays (50 Å– 0.06 Å).	Physical.
	C.T.	EM. X-rays.	Physical.
	NMR.	M. (> 1 Tesla, 0 Hz). and EM. ($\approx$ 1 MHz – 50 MHz).	Functional & Physical.
	Ultrasound.	A. (1-10 MHz).	Physical.
	Electrical impedance (Eyüboğlu *et al.*, 1989).	E. (1 KHz - 1 MHz).	Physical & Functional.
Microscopy.	Electron microscope.	P. electrons.	Physical.
	Ultrasound microscope.	A. (0.1 GHz - 1 GHz).	Physical.
	X-ray microscope.	EM. X-rays.	Physical.
Crystallography.	X-ray diffraction.	EM. X-rays.	Physical.
Nondestructive Testing.	Radiographic imaging.	EM. X-rays, γ rays.	
	Ultrasound.	A. (0.2 – 20 MHz).	
	Optical inspection.	EM. visible.	
Baggage Inspection.	X-ray absorption.	EM. X-rays.	Physical.
Object Location and Mapping.	RADAR.	EM. μwave (220 MHz – 35 GHz).	Physical.
	SONAR.	A. (100 Hz – 10 MHz).	Physical.
Geoscience and Land-Use.	RADAR.	EM. μwave.	Physical.
	Seismic imaging.	El. (10 Hz – 8 KHz).	
Meteorology.	RADAR.	EM. μwave.	Physical.

Table 1.2 Key: EM. = Electromagnetic, M. = Magnetic, E. = Electric, A. = Acoustic, P. = Particle, El. Elastic

If both the object's features and the emanations are small the image is usually obtained by inverting an expression which describes how the emanations are scattered by the object (e.g. Crystallography, electron and acoustic microscopy, electrical impedance imaging, RADAR, and SONAR).

The third factor considered when selecting the type of emanation that is to be used to probe an object is the adverse effects the emanations may have on the object. In fields such as medicine and nondestructive testing (Hull and Vernon, 1988; Blitz, 1991) (see Table 1.2) it is important to ensure that the ultrasonic and electromagnetic waves do not harm or alter the object being imaged. With any form of wave-motion a portion of the energy contained within the wave is absorbed and converted to heat by both the object and the transmission medium. The amount of heat energy that is absorbed and adequately dissipated is in part dependent on the wave's intensity and wavelength, which must therefore be appropriately selected to avoid damage. With ultrasound, pressure gradients created in the medium by the waves can also have a destructive effect (as used to destroy kidney stones), however, by appropriate selection

of waveform intensity and wavelength this can again be avoided. The use of ionising radiation (i.e., X-rays) and the adverse effects that this has on living tissue is also of major concern to those involved in medical imaging. For this reason techniques which do not use ionising radiation, such as ultrasound (Cameron and Skofronick, 1978; McCann *et al.*, 1988) and NMR imaging (Pykett, 1982; Hinshaw and Lent, 1983), are used in preference if practicable.

The final factor dictating what form the emanations take is the type of information the observer requires on the object. Table 1.2 lists a number of different imaging techniques that can provide different forms of information on objects of interest. For example: in medical imaging functional information may be required in addition to physical information and for this reason NMR imaging could be used in preference to X-ray CT; also in medical imaging a 2-D projection of an object may provide the required information and so normal X-ray absorption techniques are used in preference to X-ray CT which can provide 3-D information; in nondestructive testing the observer may be interested in flaws within an object and therefore use penetrating radiation to obtain an image in preference to visual techniques which can only provide information on the exterior of the object; and in Geoscience or landscape classification the observer may be interested in certain materials or structures and therefore chose emanations or wavelengths that allow the objects to be readily distinguished from others in the resulting image. In remote probing, having the ability to select the emanation and wavelength does not imply that all objects can be safely imaged with the required resolution. In some situations the medium will be impervious to any emanation at the wavelength required to obtain the required resolution or the object may simply be so far away that generating emanations with sufficient energy to image the object is currently impractical. Astronomical imaging is one area of imaging science where the latter applies and humankind must be content with being passive observers as there are very few celestial objects which can be imaged by remote probing techniques.

In the preceding discussion on remote probing the fact that the measurements do not always contain a direct representation of the required properties of the object was alluded to. As with most forms of scientific observation, before anything really useful can be obtained from the measurements, it is necessary to describe theoretically, in terms of mathematics, the physical process being observed. For this reason mathematical models have been developed to describe each of the image formation processes that are behind the imaging techniques in Tables 1.1 and 1.2. These models describe how the emanations interact with the transmission medium, object, emanation collection and focusing devices such as lenses, and recording devices such as films and sensor arrays (e.g. charge coupled device arrays). If the measurements do not directly contain the information required it is then usually possible to invert the expression describing the image formation process to obtain the required information. In practice it is found that the information obtained after the inversion process is of varying degrees of accuracy and depends on the formulation describing the image formation process and the accuracy of the initial measurements (Bates and McDonnell, 1986; Biemond *et al.*, 1990). Some examples of imaging techniques where inverse methods must be applied to obtain the required information are: CT where back projection or Fourier techniques are used to obtain an image of a cross-section of a body (tomograph) from measurements of the X-ray absorption through the body (Robb, 1985; Bates and McDonnell, 1986), antenna aperture imaging where a complex image of a parabolic antenna's aperture is obtained from measurements of the magnitude of the emanations in the far field in order to detect the phase shifts caused by imperfections in the surface of the antenna (Silver, 1949) and in crystallography where the magnitude of a scattered X-ray field is measured to allow the crystalline or macromolecular structure of a substance to be determined (Millane, 1990).

In all imaging situations the image obtained is not a completely accurate representation of the object's features of interest. The finite resolution of the measurements and other physical imperfections in the imaging system degrade the final image. The imperfections that commonly occur in imaging systems are: blurring of the image due to imperfections in the focusing mechanism or the relative motion between the object and observer, noise due to such phenomena as the thermally induced motion of particles and random imperfections in various components comprising the imaging system, and nonlinearities in the imaging system and recording devices such as film. As with the overall imaging process it is also possible to model, to varying degrees of accuracy, each of the above degradation processes. In any one imaging situation the degradation caused by each of the above processes will not be equal and it is therefore only usual or practical to model those degradation processes that make a significant contribution to the inaccuracies. Hence, with some prior knowledge of the degradation processes within an imaging system, it is possible to invert the expressions describing the image degradation and thereby obtain a more accurate image. In practice it is found that a model describing a degradation process in one imaging system has a very similar form to those models which describe the degradation processes in a number of other imaging systems. Thus developing a technique to overcome the degradation in one imaging system usually has far reaching effects due to its applicability to a number of imaging systems.

The ability to develop models of imaging systems is also advantageous in a number of imaging situations, such as crystallography and antenna aperture imaging, where it is not practical to measure all the information that is ideally required to produce the image. In these two situations the observer would like to measure both the magnitude and phase of the emanations, however in both cases the observer can only practically measure the magnitude. A model of the imaging process and other related mathematical relationships, however, allows the phase to be inferred from the magnitude and an image recovered (Fienup, 1982; Bates and McDonnell, 1986; Millane, 1990). It is also of interest to note that the model used to effect the retrieval of the phase is of the same form as the model used to describe image degradation due to blurring. This model is introduced and described in Section 2.3.2.

The object of the original research contained within Part One of this thesis was to develop a non-iterative algorithm to recover an improved estimate of an image from a blurred and contaminated (i.e. noisy) version of the image. This algorithm aims to invert a general expression which models how images are blurred and contaminated in a large number of different imaging situations. Thus the algorithm has not been developed to solve the blurring problem in a particular imaging system but in imaging systems in general, thereby providing benefits for a number of imaging systems. The motivation for developing this algorithm is that unlike algorithms such as inverse and Wiener filtering, which are usually used to deblur images, this algorithm does not explicitly use an estimate of the form of the blurring to obtain a solution. This algorithm uses only an estimate of the blurring to recognise the actual form of the blurring and therefore offers the promise of being less sensitive to inaccuracies in the estimates of the blurring than the more traditional techniques. This algorithm, in contrast to inverse and Wiener filtering, is known as a consistent deblurring technique as ideally it removes the actual blurring and not an estimate of the blurring. An added benefit of this new algorithm is that it is able to assist the standard phase retrieval methods to obtain a solution to the phase retrieval problem. As stated in the preceding paragraph the model describing the image blurring problem also describes the phase retrieval problem and this is the reason why the new algorithm can be used to solve both problems. It was anticipated that the new algorithm would allow the standard iterative techniques to converge to a solution faster and to overcome stagnation problems when they arise. The new algorithm has been used to deblur and retrieve the phase of a number of

contaminated images. The theory and experimental results used to develop the algorithm are contained within the remainder of Part One of this thesis along with results demonstrating the ability of the algorithm and a comparison of the new algorithm to the Wiener filter.

CHAPTER 2

PRELIMINARIES

The purpose of this chapter is to introduce the notation, terminology, and mathematical tools and relationships that are used in the remainder of this thesis. The first section of this chapter, §2.1, introduces the mathematical notation used to represent quantities such as vectors, matrices, functions, and complex numbers. Also defined are expressions which describe the level of noise or the quality of restored images. Maxwell's equations, which allow electromagnetic waves and their interactions with various media to be described mathematically, are introduced in §2.2.1. The two image formation processes, which are most relevant to the new deblurring/phase retrieval algorithm described herein, are described in detail in §2.3 along with their models. In this section the various image degradation processes, and their mathematical descriptions, are also described. The related processes of convolution and correlation, which are particularly relevant to the deblurring and phase retrieval problems, are then formally introduced in §2.4. The Fourier transform is introduced in §2.5 in both its continuous and discrete forms. The ramifications of the necessarily discrete or sampled nature of real world images are also considered in §2.5 and the continuous Fourier transform of a discrete image is derived due to its relevance to the work contained within Part One of this thesis. In §2.6 the various properties of the Fourier transform are presented and those which are required in subsequent chapters are dealt with in some detail. The Z transform and its relationship to the Fourier transform is introduced in §2.7. The properties of the Z transform are also given in §2.7. Finally, entire functions and their properties, which allow images to be completely described by the zeros of their Z or Fourier transform, are examined in §2.8.

2.1 NOTATION

Throughout this thesis vectors are denoted by bold lower case roman letters and the nth component of the vector $\mathbf{a}$ is $\mathbf{a}_n$. In a similar manner matrices are represented by a bold upper case roman letter and the individual elements of the matrix are signified by lower case subscripts. Thus, the mth component of the nth column of the matrix $\mathbf{A}$ is $\mathbf{A}_{n,m}$. The extent of a vector or matrix is indicated by the upper case form of the letter used in the subscripts, i.e. the extent of $\mathbf{a}_n$ is N,where n takes values in the range $1 \ldots N$, and $\mathbf{A}_{n,m}$ is a N by M matrix. The transpose and inverse of the matrix $\mathbf{A}$ is denoted by $\mathbf{A}^T$ and $\mathbf{A}^{-1}$ respectively.

Functions of continuous variable(s) are indicated by enclosing the variable(s) in round brackets, therefore, $f(x)$ is a function of the continuous variable x. Similarly, functions of a discrete variable(s) are conveyed by enclosing the variable(s), which can only take on integer values, in square brackets, i.e. $f[n, m]$ is a discrete function of the discrete variables n and m. When a function operates on the result of another functional operator this is expressed by

placing the first function within a set of curly brackets. Hence, the result of $g(y)$ operating on $f(x)$ is $g\{f(x)\}$. When a function is not known accurately or has been estimated this is indicated by placing a circumflex over the quantity. For example, $\hat{f}(x)$ is an estimate of the function $f(x)$.

Within this thesis it is possible for an image to exist in any one of three different domains which are known as image, Fourier, and Z space. Images which directly represent the object being imaged are in image space and are represented by functions such as $f(\mathbf{x})$. The lower case functional label, f, is used to indicate that the image is in image space and the variable $\mathbf{x}$ is a coordinate in this K-dimensional space. When the Fourier transform (see §2.5) is used to map images from image space into Fourier space this is indicated by using an upper case function label and the K-dimensional Fourier space variable $\mathbf{u}$, i.e. $F(\mathbf{u})$ is the Fourier transform or spectrum of $f(\mathbf{x})$. Likewise images that have been mapped into Z-space, such as $\mathcal{F}(\boldsymbol{\zeta})$, are denoted by an upper case calligraphic function label and the Z-space variable $\boldsymbol{\zeta}$. $\mathcal{F}(\boldsymbol{\zeta})$ is known as the Z spectrum of the image $f(\mathbf{x})$. In this thesis images exist primarily in a 2-D space and so the images in image, Fourier, and Z-space are of the form $f(x, y)$, $F(u, v)$ and $\mathcal{F}(\zeta, \gamma)$ respectively. When discussing properties of general K-dimensional signals, for simplicity the term spatial domain is used to signify the space where functions of either time or space exist. Also, the terms frequency domain and Fourier space are used interchangeably to describe the space where the Fourier transform of a spatial domain function exists. The notation $(K)\int$ is used to denote the integral of any of the above K-dimensional functions over a region of K-dimensional space.

The real and imaginary components of a complex number c have been given the notation $\Re(c)$ and $\Im(c)$ respectively. Hence

$$c = \Re(c) + i\Im(c),$$

where $i = \sqrt{-1}$. The complex number c is also expressed in its polar form

$$c = |c| \exp^{iPh(c)},$$

where $|c|$ and $Ph(c)$ are the magnitude and phase of c respectively. The magnitude is given by the expression

$$|c| = \sqrt{\Re(c)^2 + \Im(c)^2},$$

and the phase by

$$Ph(c) = \arctan \frac{\Im(c)}{\Re(c)}.$$

The complex conjugate of c is denoted by c^*, where

$$c^* = \Re(c) - i\Im(c) = |c| \exp^{-iPh(c)}.$$

As stated in Chapter 1 the recorded image, $r(\mathbf{x})$, is invariably a degraded version of the ideal image $f(\mathbf{x})$. In §2.3.2 it is stated that it is usual to model this degradation as additive contamination, i.e. $r(\mathbf{x}) = f(\mathbf{x}) + c(\mathbf{x})$, where $c(\mathbf{x})$ is the contamination term. To quantitatively describe this level of contamination on the received image the signal-to-noise ratio (SNR) is used in this thesis. The SNR is defined by the expression

$$\text{SNR} = \frac{(K)\int |f(\mathbf{x})|^2 d\mathbf{x}}{(K)\int |c(\mathbf{x})|^2 d\mathbf{x}},$$

and is expressed in dB. Unless stated otherwise, the integration is performed over the support of the integrand. It is also necessary to define a quantity that describes the success of a process designed to remove the degradation on $r(\mathbf{x})$. The quantity used herein is the quality factor, Q, which is defined by the expression

$$Q = \frac{(K)\int |f(\mathbf{x})|^2 d\mathbf{x}}{(K)\int \left|f(\mathbf{x}) - \hat{f}(\mathbf{x})\right|^2 d\mathbf{x}},$$

where $\hat{f}(\mathbf{x})$ is the estimate of the image obtained from $r(\mathbf{x})$ via a restoration algorithm. This quantity is also expressed in dB for convenience.

2.2 MAXWELL'S AND RELATED EQUATIONS

2.2.1 Maxwell's Equations

Models of imaging systems where electromagnetic emanations are used to remotely probe or sense an object can usually be derived from a set of four fundamental equations. This set of equations, namely

$$\nabla \times \mathbf{h}(\mathbf{x},t) = \mathbf{j}(\mathbf{x},t) + \frac{\partial}{\partial t}\mathbf{d}(\mathbf{x},t), \tag{2.1}$$

$$\nabla \times \mathbf{e}(\mathbf{x},t) = -\frac{\partial}{\partial t}\mathbf{b}(\mathbf{x},t), \tag{2.2}$$

$$\nabla \cdot \mathbf{d}(\mathbf{x},t) = \rho, \tag{2.3}$$

$$\nabla \cdot \mathbf{b}(\mathbf{x},t) = 0, \tag{2.4}$$

which describe the behaviour of electromagnetic emanations in any medium, are known as Maxwell's equations (Ramo *et al.*, 1984). In Maxwell's equations ∇, $\times$, and $\cdot$ denote the gradient operator $\frac{\partial}{\partial \mathbf{x}}$, the vector cross product, and vector dot product respectively. The first of these equations, Eqn.(2.1), derived from Ampère's law, describes the magnetic field, $\mathbf{h}(\mathbf{x},t)$, surrounding a charge carrying conductor in terms of the conduction current density, $\mathbf{j}(\mathbf{x},t)$, and the electric flux density $\mathbf{d}(\mathbf{x},t)$. Faraday's law, the origin of Eqn.(2.2), describes the electric field, $\mathbf{e}(\mathbf{x},t)$, being induced in a medium surrounding a changing magnetic field with a magnetic flux density $\mathbf{b}(\mathbf{x},t)$. The third equation, Eqn.(2.3), is an expression of Gauss's law and states that the net flow of electric flux through a surface is equal to the electric charge density, ρ, of any charge enclosed by the surface. The final equation, Eqn.(2.4), is the magnetic equivalent of the third equation and states that the net magnetic flux flowing through any surface is zero. The reason the net flux is zero is because there is no evidence, at present, to indicate that isolated magnetic sources or sinks exist (Ramo *et al.*, 1984).

In addition to Maxwell's equations there are also three auxiliary equations which describe the fields in terms of their flux densities and the constitutive properties of the medium. The three auxiliary equations are

$$\mathbf{d}(\mathbf{x},t) = \epsilon \mathbf{e}(\mathbf{x},t), \tag{2.5}$$

$$\mathbf{b}(\mathbf{x},t) = \mu \mathbf{h}(\mathbf{x},t), \tag{2.6}$$

$$\mathbf{j}(\mathbf{x},t) = \sigma \mathbf{e}(\mathbf{x},t), \tag{2.7}$$

where ϵ, μ, and σ are the permittivity (dielectric constant), permeability, and conductivity of the medium. The permittivity and permability of free space, ϵ_0 and μ_0 respectively, are related

to the constitutive properties of another medium via the relative permittivity and permability, ϵ_r and μ_r where

$$\epsilon = \epsilon_r \epsilon_0, \tag{2.8}$$

$$\mu = \mu_r \mu_0, \tag{2.9}$$

and the refractive index, ν, of the medium is given by

$$\nu = \sqrt{\epsilon_r \mu_r}. \tag{2.10}$$

2.2.2 Wave Equations

An expression which describes the motion of an electric or magnetic field through a medium can be obtained from Maxwell's equations. The development of the expression describing the motion is the same for either field and for this reason only the electric field is considered here. It is assumed that the medium is homogeneous (i.e. ϵ, μ, and σ are constant with respect to $\mathbf{x}$) and that there are no free charges or currents within the medium (i.e. $\rho = 0$ and $\mathbf{j}(\mathbf{x}, t) = 0$). Hence,

$$\nabla \times \mathbf{h}(\mathbf{x}, t) = \frac{\partial}{\partial t}\mathbf{d}(\mathbf{x}, t) \qquad \text{and}$$

$$\nabla \cdot \mathbf{d}(\mathbf{x}, t) = 0.$$

Taking the curl of Eqn.(2.2), using the vector identity $\nabla \times \nabla \times \mathbf{a} = \nabla(\nabla \cdot \mathbf{a}) - \nabla^2 \mathbf{a}$, changing the order of the derivatives on the RHS, and substituting in Eqn.(2.6) gives

$$\nabla^2 \mathbf{e}(\mathbf{x}, t) = \mu \frac{\partial}{\partial t} \nabla \times \mathbf{h}(\mathbf{x}, t). \tag{2.11}$$

On substituting Eqns. (2.1) and (2.5) into the RHS of Eqn.(2.11) the time domain wave equation

$$\nabla^2 \mathbf{e}(\mathbf{x}, t) = \mu\epsilon \frac{\partial^2}{\partial t^2} \mathbf{e}(\mathbf{x}, t), \tag{2.12}$$

is obtained. On taking the temporal Fourier transform of both sides (see §2.5) and defining the free space wave number as $k_0 = \omega\sqrt{\epsilon_0 \mu_0}$, where $\frac{\partial}{\partial t} = i\omega$, the frequency domain wave equation, which is also known as the Helmholtz equation,

$$\nabla^2 \mathbf{E}(\mathbf{x}, k) = -\nu^2 k_0^2 \mathbf{E}(\mathbf{x}, k), \tag{2.13}$$

is obtained. The solutions to Eqns. (2.12) and (2.13) are wavelike in nature and it is for this reason that the equations are known as wave equations. For example a solution to Eqn.(2.13) is

$$\mathbf{E}(\mathbf{x}, k) = A \exp^{-i\nu \mathbf{k} \cdot \mathbf{x}}, \tag{2.14}$$

which is a plane wave, with amplitude A, propagating in the direction of the vector $\mathbf{k}$.

2.2.3 Kirchhoff-Helmholtz Formula

In this section the Kirchhoff-Helmholtz formula, which relates the field at a point within a volume containing sources to the fields incident on the circumscribing surface, is derived. The Kirchhoff-Helmholtz formula is developed here as it is a prerequisite to an expression in

§2.3.1.2 which describes the image of a microwave antenna's aperture in terms of its radiation pattern. Consider an arbitrary homogeneous region of space which has a volume V and is bounded by the surface S (see Fig. 2.1). In this volume there exists an electromagnetic field Ψ which satisfies the Helmholtz equation (Eqn.(2.13)) and is generated by sources inside V. An expression for the field at any point P_1 within the volume can be obtained with the aid of Green's theorem (Goodman, 1968; Kreyszig, 1983). Green's theorem states that if there are two complex valued functions $G(\mathbf{x})$ and $U(\mathbf{x})$, which are continuous and have continuous first and second order derivatives within and on the surface S, surrounding a volume V, then

$$\int_V \left(G\nabla^2 U - U\nabla^2 G\right) dv = \int_S \left(G\frac{\partial U}{\partial n} - U\frac{\partial G}{\partial n}\right) ds \tag{2.15}$$

In Eqn.(2.15) it is conventional for the unit normal $\mathbf{n}$ to be directed out of the volume V as indicated in Fig. 2.1.

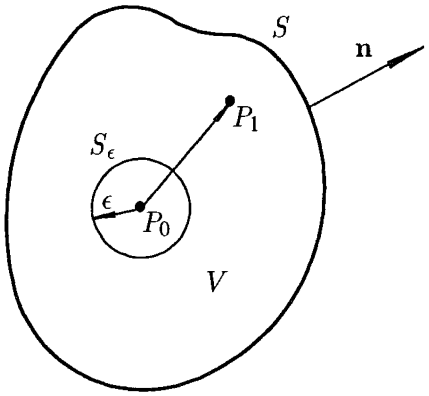

Figure 2.1 The surfaces of integration, S and S_ϵ, used to derive the Kirchhoff-Helmholtz formula.

It is useful at this point to define a Green's function,

$$g(P_0, P_1) = \frac{\exp^{ikr_{01}}}{r_{01}}, \tag{2.16}$$

where r_{01} is the length of the vector $\mathbf{r}_{01}$ pointing from P_0 to P_1 in Fig. 2.1. The Green's function chosen satisfies Helmholtz's equation (Eqn.(2.13)) in V and describes a unit amplitude spherical wave expanding from a source at P_0. On letting $U = \Psi$ and $G = g$, Eqn.(2.15) reduces to

$$\int_S \left(g\frac{\partial \Psi}{\partial n} - \Psi\frac{\partial g}{\partial n}\right) ds + \int_{S_\epsilon} \left(g\frac{\partial \Psi}{\partial n} - \Psi\frac{\partial g}{\partial n}\right) ds = 0, \tag{2.17}$$

where the singularity at P_0, arising from the Green's function tending to infinite value, has been excluded by enclosing P_0 by the surface S_ϵ. An expression for the field at P_1 can now be obtained by evaluating the integral over S_ϵ as ϵ tends to zero. On S_ϵ, $r_{01} = \epsilon$ and so the integration over S_ϵ becomes

$$\int_{S_\epsilon} \left\{\frac{\exp^{ik\epsilon}}{\epsilon}\frac{\partial}{\partial n}\Psi(r_{01})\Big|_{r_{01}=\epsilon} - \Psi(r_{01})\Big|_{r_{01}=\epsilon}\left[\frac{\exp^{ik\epsilon}}{\epsilon^2} - ik\frac{\exp^{ik\epsilon}}{\epsilon}\right]\right\} ds.$$

Now letting $\epsilon \to 0$

$$\begin{aligned}\int_{S_\epsilon} \left(g\frac{\partial \Psi}{\partial n} - \Psi\frac{\partial g}{\partial n}\right) ds &= 4\pi\epsilon^2 \left\{\frac{\exp^{ik\epsilon}}{\epsilon}\frac{\partial}{\partial n}\Psi(P_0) - \Psi(P_0)\left[\frac{\exp^{ik\epsilon}}{\epsilon^2} - ik\frac{\exp^{ik\epsilon}}{\epsilon}\right]\right\} \\ \lim_{\epsilon\to 0} &= -4\pi\Psi(P_0),\end{aligned}$$

and therefore Eqn.(2.17) becomes

$$\Psi(P_0) = \frac{1}{4\pi}\int_S \left(g\frac{\partial \Psi}{\partial n} - \Psi\frac{\partial g}{\partial n}\right) ds, \tag{2.18}$$

which is the Kirchhoff-Helmholtz formulation.

2.3 IMAGE FORMATION

2.3.1 Ideal Images

In this section two different imaging processes, optical imaging and electromagnetic imaging of antenna apertures, are described along with the mathematical models, derived from electromagnetic theory, used to describe each of the image formation processes. The two image formation processes are described in detail since the new deblurring algorithm can be applied directly to images obtained from the former process, whilst the phase retrieval version of the algorithm can be used to solve the phase problem in the latter process. It is shown that the optical image formation process can be described in terms of what is known as geometrical optics. This in turn allows the blurring process to be straightforwardly described and modelled without unnecessarily complicating the treatment.

2.3.1.1 Optical Image Formation

Imaging systems that use electromagnetic energy at visible or near visible wavelengths (hereafter called light) and optical imaging techniques are used in the vast majority of imaging situations. A schematic of a typical optical imaging system is given in Fig. 2.2. In this system light from a natural or artificial source (e.g. the sun or an incandescent lamp) is reflected by the 3-D object in object space whose image is to be obtained. A lens system collects a portion of this reflected light and directs it onto the imaging plane in image space where an image is formed. The system which collects the light and focuses it onto the imaging plane is known herein as the image formation system. In order to record this image a 2-D light sensitive recording device, such as photographic film or an array of photosensitive devices, can be positioned on the imaging plane and this device is exposed to the incident light for an appropriate time interval known as the exposure. The system which records the image formed on the imaging plane is termed the image recording system. In an ideal imaging system all the collected light from a single point on the object would be focused onto a single point on the imaging plane forming the ideal image. In practice it is not possible to design a lens system that can map every point in 3-D object space onto a single corresponding point on the imaging plane and the consequences of this are discussed in §2.3.2.

The ideal image, represented by the function $f(x, y)$, has both finite support (i.e. non-zero widths L_x and L_y in the x and y directions respectively) and finite energy since the object and/or the recording device has a finite extent and a necessarily finite amount of energy is absorbed by the recording device. Images which have both finite support and finite energy are known as compact images (Fright, 1984). In optical imaging situations the function $f(x, y)$ is a real function which cannot take on negative values since the smallest amount of energy that can be absorbed by the recording device is zero (Andrews and Hunt, 1977). Images which are real and cannot take on negative values are known as positive images (Bates and McDonnell, 1986). Note however that interferometric techniques, such as holography, can be used to modify the nature of the information recorded in the image.

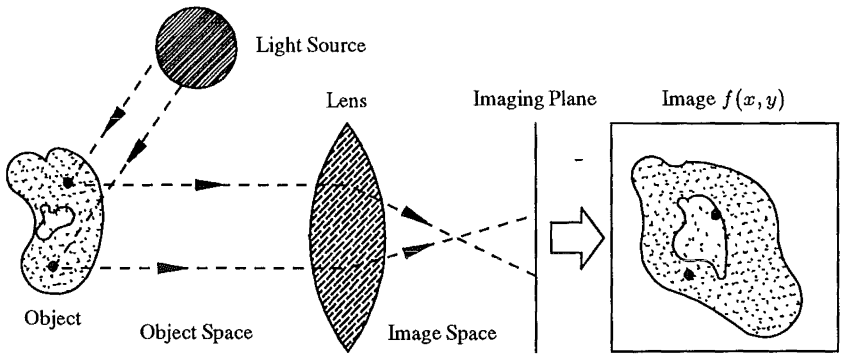

Figure 2.2 Schematic of an ideal optical imaging system

In imaging situations where the wavelength of the electromagnetic waves is much smaller than the changes in the constitutive properties of the media it is appropriate to model the wave motion, within those regions where there is a negligible change in the constitutive properties, with a plane wave (Born and Wolf, 1970). The imaging techniques in Tables 1.1 and 1.2 that can be modelled by using this *high frequency* approximation are the optical imaging techniques (e.g. archival imaging, optical satellite imaging, microscopy, and optical astronomy), X-ray absorption imaging techniques, radar, and sonar. In the local regions where the medium is homogeneous the plane wave is of the same form as that described by Eqn.(2.14), i.e.

$$P(\mathbf{x}, k) = a \exp^{-i\mathbf{k}\cdot\mathbf{x}}, \tag{2.19}$$

where a is a constant which may take on complex values (Born and Wolf, 1970). The surfaces in space on which the wave's phase

$$\mathbf{k} \cdot \mathbf{x} = \text{constant},$$

are known as wavefronts and these surface are perpendicular to the direction of the wave's motion. Thus, by determining where the wavefronts lie in space (a trivial operation in a homogeneous medium), it is possible to plot with *rays* the path of propagation through the medium.

In an inhomogeneous medium the wave's velocity alters with the changes in the constitutive properties of the media which in turn changes the orientation of the wavefronts and the direction of propagation. In a medium of this type, where the high frequency approximation still applies, it is possible to describe the wave motion using a plane wave of the form

$$P(\mathbf{x}, k) = a(\mathbf{x}) \exp^{-ik_0 h(\mathbf{x})}, \tag{2.20}$$

where $a(\mathbf{x})$ and $h(\mathbf{x})$ are real functions of position in the medium (Jones, 1986a; Born and Wolf, 1970). The function $h(\mathbf{x})$ describes the *optical path* or direction of propagation in the medium (Born and Wolf, 1970). The manner in which this wave propagates through the medium can then be found from the position of the wave's wavefronts by establishing where

$$h(\mathbf{x}) = \text{constant}.$$

To determine the form of $h(\mathbf{x})$ Eqn.(2.20) is substituted into the Helmholtz equation (Eqn.(2.13)), where the refractive index ν is now a function of $\mathbf{x}$, and the same powers of k_0

are grouped together giving (Born and Wolf, 1970; Jones, 1986a; Bates *et al.*, 1991)

$$\begin{aligned} & k_0{}^2 \left[\left(\nu(\mathbf{x})^2 - \nabla h(\mathbf{x}) \cdot \nabla h(\mathbf{x}) \right) a(\mathbf{x}) \right] \\ -ik_0 & \left[2\nabla a(\mathbf{x}) \cdot \nabla h(\mathbf{x}) + a(\mathbf{x}) \nabla^2 h(\mathbf{x}) \right] = -\nabla^2 a(\mathbf{x}). \end{aligned} \tag{2.21}$$

As the high frequency approximation used to derive Eqn.(2.21) relies on the fact that the wavelength is small in comparison to the changes in the constitutive properties, a significant change in $a(\mathbf{x})$ occupies many wavelengths and therefore $\nabla a(\mathbf{x})$ is very small and k_0 is very large in comparison. On removing the insignificant terms from Eqn.(2.21), the Eiokonal equation of geometric optics is formed,

$$\nabla h(\mathbf{x}) \cdot \nabla h(\mathbf{x}) = \nu(\mathbf{x})^2, \tag{2.22}$$

which describes the phase of the plane wave in terms of the refractive index (Born and Wolf, 1970; Jones, 1986a; Bates *et al.*, 1991). On equating the second most significant term in Eqn.(2.21) to zero the first transport equation

$$2\nabla a(\mathbf{x}) \cdot \nabla h(\mathbf{x}) + a(\mathbf{x}) \nabla^2 h(\mathbf{x}) = 0, \tag{2.23}$$

which can be used to determine the amplitude of the plane wave at $\mathbf{x}$, is obtained. Thus, by solving Eqns. (2.22) and (2.23) it is possible to determine the amplitude and path of the plane wave as it propagates through the medium. It is possible to obtain an exact solution to Eqn.(2.22), however, in general Eqn.(2.23) cannot be solved explicitly and so approximate techniques such as ray tracing, geometrical optics, and the geometrical theory of diffraction are used to obtain solutions to these equations (Born and Wolf, 1970; Enright, 1992).

When modelling the optical image formation process it is not necessary to solve the Eiokonal and first transport equations, since geometrical optics, the laws of which can be derived from the Eiokonal equation, provide a perfectly adequate description of the imaging process. It is also of interest to note that the description of optical systems in terms of the geometrical relationships of geometrical optics was occurring long before the theory of electromagnetics was formally understood. The laws of geometric optics which describe image formation in terms of straight rays (as opposed to curved rays (Enright, 1992)), the law of reflection, and the law of refraction (Snell's law) are derived from two main assumptions. The first assumption, as with the derivation of the Eiokonal equation, is that the wave's wavelength is much smaller than the changes in the medium's refractive index and it is therefore appropriate to model the wave motion by a plane wave. The second assumption is that the changes in the refractive index occur at a relatively small number of well defined boundaries. The result of these assumptions is that in the large regions where the medium is homogeneous (implied by the second assumption) it is possible to plot the propagation of the waves with straight rays, since Eqn.(2.19) adequately describes the wave motion. At the refractive index boundaries the laws of reflection and refraction, which are derived below, can then be used to determine how the direction of the rays change at the boundary.

The laws of reflection and refraction are derived by considering the effect of a refractive index boundary on the plane wave

$$P_i(\mathbf{x}, k) = a_i \exp^{-i\nu_1 k_0 \alpha}, \tag{2.24}$$

which is propagating in the α direction towards the boundary in Fig. 2.3. At the boundary a portion of the wave's energy is reflected, represented by the reflected wave

$$P_r(\mathbf{x}, k) = a_r \exp^{i\nu_1 k_0 \alpha'}, \tag{2.25}$$

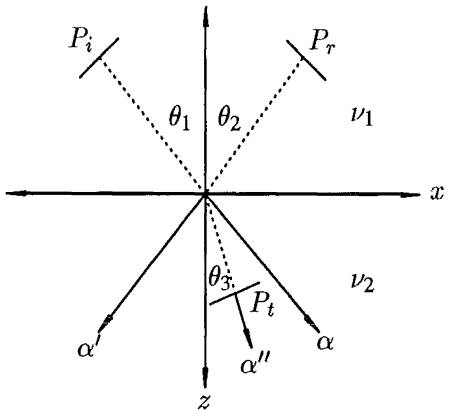

Figure 2.3 Schematic showing the interaction of a plane wave P_i with a boundary that separates two media of different refractive index ν_1 and ν_2. P_r and P_t are the reflected and refracted waves respectively.

whilst the remainder, represented by the wave

$$P_t(\mathbf{x}, k) = a_t \exp^{-i\nu_2 k_0 \alpha''}, \tag{2.26}$$

undergoes refraction at the boundary. On using the coordinate transformations

$$\begin{aligned} \alpha &= x \sin\theta_1 + z\cos\theta_1 \\ \alpha' &= -x\sin\theta_2 + z\cos\theta_2 \\ \alpha'' &= x\sin\theta_3 + z\cos\theta_3, \end{aligned}$$

setting $x = 0$, and equating the phases of Eqns. (2.24), (2.25), and (2.26) (since phase must be continuous on the boundary) it is found that the phases are related by the expression

$$\nu_1 \sin\theta_1 = \nu_1 \sin\theta_2 = \nu_2 \sin\theta_3.$$

Thus the angle at which a wave is reflected from a surface is equal to the angle at which the boundary is approached and the angle of refraction, θ_3, is given by

$$\sin\theta_3 = \frac{\nu_1}{\nu_2}\sin\theta_1 \tag{2.27}$$

The laws of geometric optics, and the arguments used above to obtain these laws, can be used to obtain equations which describe the imaging process in imaging systems containing lenses. To obtain these equations it is necessary to consider the refraction of a plane wave as it passes across a single spherical boundary separating two different media (e.g. if b_1 was the only refractive index boundary in Fig. 2.4). On analysing how the ray is reflected from this spherical refractive surface, which has a radius of curvature R_1, it is found that the object's position along the lens axis (see Fig. 2.4), s, is related to the images position, s', by the relationship

$$\frac{\nu_1}{s} + \frac{\nu_2}{s'} = \frac{\nu_2 - \nu_1}{R_1},$$

where ν_1 and ν_2 are the refractive indices of the medium in which the rays originate and terminate respectively (Sears *et al.*, 1980).

For a lens, which has a refractive index ν_2 and has been placed in a medium with a refractive index ν_1, refraction from the two spherical surfaces b_1 and b_2 in Fig. 2.4 must be considered.

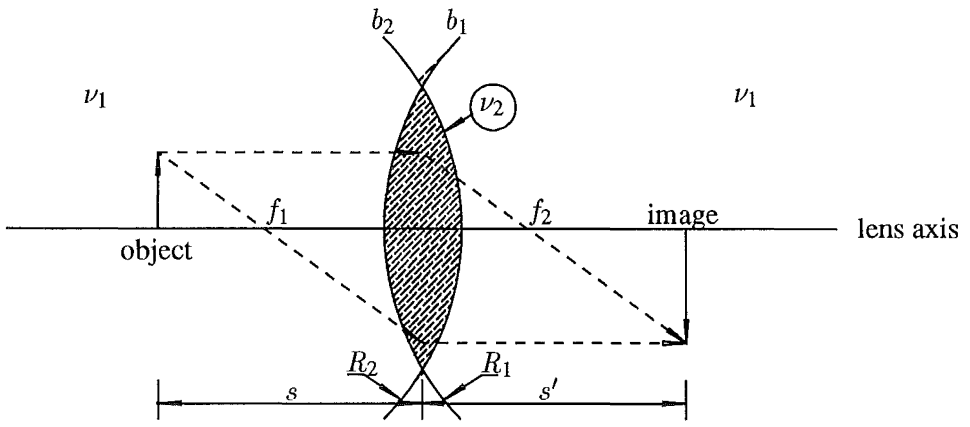

Figure 2.4 The formation of an image, at s', of an object situated at s by an ideal thin lens.

On neglecting the thickness of the lens and using the knowledge that the rays incident on the second boundary originate from the image formed by the first, the equation

$$\frac{\nu_1}{s} + \frac{\nu_1}{s'} = (\nu_2 - \nu_1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right), \tag{2.28}$$

where R_2 is the radius of curvature of b_2, is obtained. The primary focal point, f_1, and secondary focal length, f_2, of a lens are defined as the point on the lens axis onto which all the rays from an object an infinite distance from the lens are focused and the position of a point object on the lens axis which forms an image at an infinite distance along the lens axis, respectively. Setting $s' = \infty$ gives the lensmaker's equation

$$\frac{\nu_1}{f_1} = (\nu_2 - \nu_1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right), \tag{2.29}$$

which describes the position of the second focal point of the lens. On substituting Eqn.(2.29) into Eqn.(2.28) the familiar thin lens equation

$$\frac{1}{f_1} = \frac{1}{s} + \frac{1}{s'}, \tag{2.30}$$

is obtained (Sears *et al.*, 1980). Using the above equations it is now possible to describe how every point on an object is mapped into image space by an imaging system in which the thin lens and high frequency approximations apply.

2.3.1.2 Antenna Aperture Imaging

Antennae form an integral part of a large number of remote probing, remote sensing, and communication systems where the radiation and/or reception of electromagnetic waves is required. When designing or selecting an antenna for use in a particular application it is necessary to determine how well the requirements of the application are met by the properties of the various antennae. The antenna properties that must be considered are: the efficiency with which the antenna transfers the electromagnetic energy between the transceiver and the medium, the bandwidth of the electromagnetic waves, the polarisation of the waves, and the antenna's radiation pattern (Conner, 1989). The radiation pattern is the property that dictates

to the largest degree what size, shape, and type of antenna is used for in a particular situation. For example, antennae for RADAR systems are required to form radiation patterns where the majority of the radiated energy is confined to a narrow main beam. The reason for this requirement is that the resolution of a RADAR system is directly related to the width of the main beam. In contrast, an antenna for a base station which broadcasts to a number of sites spread out over a wide area should ideally radiate equally in all directions azimuthally. At microwave frequencies the antenna designer is in the advantageous position that *aperture radiators*, such as reflectors and lenses, can be used to modify a primary radiation pattern, emitted by a primary antenna or feed, to produce a secondary radiation pattern which meets the requirements of the application (Silver, 1949; Ramo *et al.*, 1984; Conner, 1989). As the wavelengths of microwaves are generally much smaller than the dimensions of a microwave antenna it is possible to use the laws of geometric optics and the geometric optics approximations (see §2.3.1.1) to describe the form of the radiation pattern emitted by microwave antennae (Silver, 1949; Conner, 1989). The remainder of this section is concerned with deriving an expression which describes the form of the secondary radiation pattern of an aperture antenna, at various distances from the aperture, in terms of the shape and size of the aperture. Being able to derive the shape of the aperture from measurements of the secondary field is of immense importance as: aperture antennae are in extensive use, and any defects in the surface of the reflector or in the lens distorts the secondary radiation pattern or does not allow the received radiation to be correctly focused onto the primary antenna.

Before deriving an expression for the secondary field it is first necessary to derive a schematic describing the antenna's reflector or lens in terms of an aperture through which the electromagnetic radiation passes. Consider initially the aperture antenna in Fig. 2.5(a), where the reflective surface S_R is being illuminated by the electromagnetic field ψ originating from the point source P_S. Using the principles of geometric optics it is possible to draw a ray diagram, represented by the dashed and thin solid lines in Fig. 2.5(a), describing the path of the incident and reflected electromagnetic waves. When the reflected rays lie entirely within one hemisphere of space and a curve C can circumscribe all the reflected rays, then it is possible to think of the field as if it was created by the equivalent source P'_S to the left of the antenna (see Fig. 2.5(a)) (Silver, 1949). The surface S_R can therefore be thought of as an aperture through which all the energy passes. In the remainder of this section, for simplicity, it is assumed that the aperture, A, is a hole in an opaque 2-D surface of infinite extent, S_1, and the antenna can be represented by the schematic in Fig. 2.5(b). It is obvious that the schematic for the aperture antenna formed with a lens has exactly the same form as that in Fig. 2.5(b) since in this case the actual source is P'_s.

In the schematic in Fig. 2.5(b) $\mathbf{n_a}$ is a vector normal to the aperture surface, $P(x, y, z)$ is the point from where the secondary field is to be observed, $\mathbf{r}_0$ is the unit vector originating at the origin and heading in the direction of P, and if $(\xi, \eta, 0)$ is the position of a point on the aperture surface then $\mathbf{r}_1$ is a unit vector originating at (ξ, η) and also heading in the direction of P. For the purposes of the expression to be developed here, describing the secondary field, it is assumed that the distance between the aperture and P, r_1, is much greater than the dimensions of the aperture and is equal to many wavelengths. The fields on the surfaces A and S_1 are defined by what is known as Kirchhoff's boundary conditions (Goodman, 1968). These boundary conditions state that:

1. The field ψ and its derivative $\frac{\partial \psi}{\partial n_a}$ on A are the same as if the opaque surface S_1 was not present.

2. The field ψ and its derivative $\frac{\partial \psi}{\partial n_a}$ are identically zero over the whole of the opaque

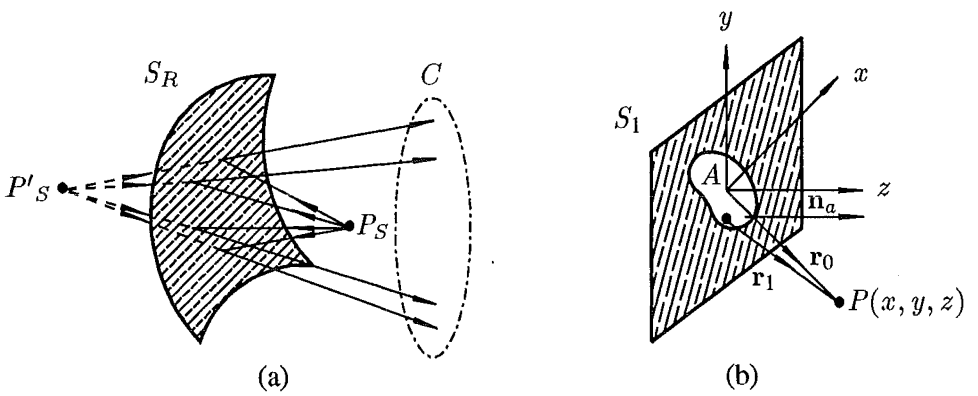

Figure 2.5 (a) Ray diagram showing the formation of the secondary field of a reflector antenna, (b) a schematic that can be used to represent the antenna in (a).

surface S_1.

In practice the field does not completely obey the boundary conditions, however, any departures from these conditions have a very small effect on the secondary field, at the distances from the aperture to be considered here, and can therefore be neglected (Goodman, 1968).

The Kirchhoff diffraction formula, which describes the form of an arbitrary field after it has passed through an aperture in an opaque surface of infinite extent (see Fig. 2.5), can now be obtained from the Kirchhoff-Helmholtz formula (Eqn.(2.18)). The hemispherical region of space to the right of S_1 in Fig. 2.5 can be thought of as being bounded by the three surfaces S_1, A, and S_∞ where S_∞ is a hemispherical surface of radius $r_0 \to \infty$ which caps this hemispherical region of space. Using Eqn.(2.18) the field at any point to the right of the aperture is given by

$$\Psi(P_0) = \frac{1}{4\pi} \int_{S_1 \cup A \cup S_\infty} \left(g \frac{\partial \Psi}{\partial n} - \Psi \frac{\partial g}{\partial n} \right) ds, \tag{2.31}$$

As stated earlier in this section the fields on S_1 and A satisfy the Kirchhoff boundary conditions and therefore the contribution to Eqn.(2.31) from the integration over S_1 is zero. To determine what contribution the integration over S_∞ makes to the value of $\Psi(P_0)$ it is necessary to determine how the energy of a wave propagating away from an aperture changes if the energy of the field is to be conserved. Since the surface S_∞ increases in area at a rate equal to ${r_0}^2$ the field in this region must diminish as rapidly as $1/r_0$ if the field is to satisfy the requirements of the conservation of energy (Elliot, 1966). Fields that vanish at least as fast as a diverging spherical wave (i.e. at least as fast as $1/r_0$) satisfy the Sommerfield radiation condition

$$\lim_{R \to 0} r_0 \left(\frac{\partial \Psi}{\partial n} - ik\Psi \right) = 0,$$

consequently, the integration over S_∞ contributes nothing to $\Psi(P_0)$ (Goodman, 1968). The field to the right of the aperture is therefore given by the expression

$$\Psi(P_0) = \frac{1}{4\pi} \int_A \left(g \frac{\partial \Psi}{\partial n} - \Psi \frac{\partial g}{\partial n} \right) ds, \tag{2.32}$$

which is Kirchhoff's diffraction formula.

An expression for the secondary field of an aperture antenna can now be obtained from Kirchhoff's diffraction formula by using the high frequency approximation, which allows the field in the region of the aperture to be expressed in the form of the plane wave

$$\Psi(\xi,\eta) = a(\xi,\eta)\exp^{-ik_0 h(\xi,\eta)},$$

where $h(\xi,\eta)$ = constant and $a(\xi,\eta)$ define the wavefronts and amplitude of the field respectively (see §2.3.1.1). Using the above expression for the aperture field, the derivative of the aperture field with respect to the outward facing normal, $\mathbf{n}$, is given by

$$\begin{aligned}\frac{\partial\Psi}{\partial n} &= -\mathbf{n}_a\cdot\nabla\Psi \\ &= ik_0\Psi\mathbf{n}_a\cdot\nabla h - \frac{\Psi}{a}\frac{\partial a}{\partial n}\end{aligned} \tag{2.33}$$

The high frequency approximation implies that k_0 is very large and therefore the second term in Eqn.(2.33) is negligible in comparison to the first. On defining a unit vector $\mathbf{s}$, which originates at a point on the aperture and points in the direction of propagation of the field, it is then possible, using the Eiokonal equation (Eqn.(2.22)), to express ∇h in the form

$$\nabla h = \nu\mathbf{s} = \frac{k}{k_0}\mathbf{s}.$$

It is therefore possible to express Eqn.(2.33) in the form

$$\frac{\partial\Psi}{\partial n} = ik\Psi\mathbf{n}_a\cdot\mathbf{s}. \tag{2.34}$$

Similarly, the derivative of the Green's function with respect to the outward facing normal is given by

$$\begin{aligned}\frac{\partial\Psi}{\partial n} &= -\mathbf{n}_a\cdot\nabla g \\ &= -\left(ik-\frac{1}{r_1}\right)g\mathbf{n}_a\cdot\mathbf{r_1}\end{aligned} \tag{2.35}$$

On substituting Eqns. (2.34) and (2.35) into the Kirchhoff diffraction formula (Eqn.(2.32)) the expression

$$\Psi(P_0) = \frac{1}{4\pi}\int_A \frac{\exp^{ikr_1}}{r_1}a(\xi,\eta)\exp^{-ik_0h(\xi,\eta)}\left[\left(ik-\frac{1}{r_1}\right)\mathbf{n}_a\cdot\mathbf{r_1} + ik\mathbf{n}_a\cdot\mathbf{s}\right]ds, \tag{2.36}$$

which describes the secondary field of an aperture antenna, is obtained.

At distances equivalent to many wavelengths from the aperture Eqn.(2.36) can be simplified by making two sets of assumptions. The first set of assumptions applies in a region of space, to the right of the aperture, known as the Fresnel region. The Fresnel region is defined to be the region of space where R_1 is much greater than both the wavelength and the values of ξ and η. The relative difference between the sizes of the above quantities allows the following approximations to be made in the Fresnel region (Silver, 1949):

1. As $r_1 >> k$ the $1/r_1$ term in round brackets in Eqn.(2.36) is much smaller than k and can be neglected.

2. At large distances from the aperture $\mathbf{r_1}$ can be considered to be parallel to $\mathbf{n}_a$ and therefore $\mathbf{n}_a\cdot\mathbf{r_1}$ can be replaced by unity.

3. At large distances $r_1 \approx z$ and the $1/r_1$ term outside the square brackets can be replaced by $1/z$.

4. An approximation to the phase term $\exp^{ikr_1}$ can also be made, however, since the difference between r_1 and r_0 could be significantly different in terms of the phase of the wave, this approximation must be carried out more carefully than those above. As

$$r_1 = \sqrt{z^2 + (x-\xi)^2 + (y-\eta)^2}$$

the binomial expansion can be used to express r_1 in the form

$$r_1 = z + \frac{(x-\xi)^2 + (y-\eta)^2}{2z},$$

where terms in the expansion higher than the second order have been neglected. Thus, the phase term can be expressed in the form

$$\exp^{ikz}\exp^{\frac{ik}{2z}(x^2+y^2)}\exp^{\frac{ik}{2z}(\xi^2+\eta^2)}\exp^{\frac{-ik}{z}(x\xi+y\eta)}.$$

On substituting all the approximations into Eqn.(2.36), and using the relationship $k = 2\pi/\lambda$, the expression

$$\begin{aligned}\Psi(x,y) &= \frac{1}{2\lambda z}\exp^{ikz}\exp^{\frac{ik}{2z}(x^2+y^2)}\int_A a(\xi,\eta)\exp^{-ik_0h(\xi,\eta)} \\ &\quad \exp^{\frac{ik}{2z}(\xi^2+\eta^2)}\exp^{\frac{-ik}{z}(x\xi+y\eta)}\,[1+\mathbf{n}_a\cdot\mathbf{s}]\,d\xi d\eta, \qquad (2.37)\end{aligned}$$

describing the field an observer situated in the Fresnel region would observe, is obtained.

At distances from the antenna's aperture where $z >> \frac{k}{2}(\xi^2+\eta^2)$ an observer would be situated in what is known as the Fraunhofer or far field region of the antenna's field (Silver, 1949). In the far field the phase term

$$\exp^{\frac{ik}{2z}(\xi^2+\eta^2)} = 1$$

and the field an observer situated in this region would observe is given by the expression

$$\Psi(x,y) = \frac{1}{2\lambda z}\exp^{ikz}\exp^{\frac{ik}{2z}(x^2+y^2)}\int_A a(\xi,\eta)\exp^{-ik_0h(\xi,\eta)}\exp^{\frac{-ik}{z}(x\xi+y\eta)}\,[1+\mathbf{n}_a\cdot\mathbf{s}]\,d\xi d\eta.$$

Furthermore, if it is assumed that the phase on the aperture is constant (i.e. the phase errors are small) then $\mathbf{n}_a \cdot \mathbf{s} = 1$ (Silver, 1949) and the above equation can be simplified to give the more familiar expression for the field,

$$\Psi(x,y) = \frac{1}{\lambda z}\exp^{ikz}\exp^{\frac{ik}{2z}(x^2+y^2)}\int_A a(\xi,\eta)\exp^{-ik_0h(\xi,\eta)}\exp^{\frac{-ik}{z}(x\xi+y\eta)}d\xi d\eta. \qquad (2.38)$$

The significance of Eqn.(2.38) becomes clear if it is compared to the definition of the Fourier transform in §2.5. It can be seen from this comparison that the far field of an aperture antenna is simply the Fourier transform of its aperture field scaled by a quadratic phase term. The upshot of this relationship is that by inverse Fourier transforming the measured far field it is possible to obtain an image of the antenna's aperture and thereby determine the nature of any lens defects. In practice it is only feasible to measure the magnitude of the far field and for this reason the phase of the far field must be recovered before an image of the aperture can be obtained. The principles of the phase retrieval problem are discussed in §3.5 and the results of the application of the new algorithm to this problem are described in §6.2.1.

2.3.2 Image Degradation

As alluded to in Chapter 1 and §2.3.1.1 all images formed by practical imaging systems are degraded to some extent and the recorded image $g(x, y)$ is therefore a degraded version of the ideal image $f(x, y)$. There are three main types of degradation that can corrupt $f(x, y)$ and these are categorised as either blurring, nonlinear effects or noise (Andrews and Hunt, 1977; Biemond *et al.*, 1990). The origin, effects and models used to describe each of these forms of degradation are described in the following paragraphs and an overall model for image degradation is derived from this discussion.

The forms of degradation that can be grouped together under the category of blurring all arise from the nonideal way in which the images are formed in practical imaging systems. If it is assumed that the image formation system is a linear time invariant (LTI) system (see §2.4 for a description of LTI systems), with an input and output which are $f(x, y)$ and the blurred image $b(x, y)$ respectively, it is possible to develop a model of the blurring process (Andrews and Hunt, 1977). To obtain this model consider the output from such a LTI system when the input is an impulse situated at the point (s, t) on the image plane. In this situation the output is simply the impulse response $h(x, y, s, t)$ of the image formation system. If this impulse is now weighted by the value of the ideal image at (s, t) then the output of the system would be $h(x, y, s, t)f(s, t)$. When the input to the system is $f(x, y)$, which can be represented by an infinite number of impulses covering the image plane and all weighted by $f(x, y)$ (see §2.4), then $b(x, y)$ is given by the expression

$$b(x, y) = \int f(s, t)h(x, y, s, t)dsdt, \tag{2.39}$$

where the integration is over the extent of the image plane. In imaging literature the imaging system's impulse response is more commonly known as the point spread function (psf), and its Fourier transform as the optical transfer function (OTF), for obvious reasons. In the situation where the psf has the same functional form for all values of (s, t) on the image plane the psf is said to be point spread invariant or isoplanatic (Andrews and Hunt, 1977; Bates and McDonnell, 1986). When the blurring is isoplanatic the psf becomes a function of only the argument differences $(x-s)$ and $(y-t)$ and Eqn.(2.39) reduces to the more familiar convolution integral

$$\begin{aligned} b(x, y) &= \int f(s, t)h(x - s, y - t)dsdt, \\ &= h(x, y) \odot f(x, y) \end{aligned} \tag{2.40}$$

where the operator $\odot$ denotes convolution (see §2.4).

In order to gain an appreciation of the form of the distortion caused by blurring and the origins of the functions commonly used to describe psfs; three of the more common types of blurring are introduced in the context of optical imaging. The first type of blurring to be described is out-of-focus blur. As alluded to in §2.3.1.1 it is not always possible to focus all the light reflected from a point in object space onto a single point on the 2-D imaging plane. The form of this distortion can be determined by considering the image formed by the imaging system in Fig. 2.6. In this imaging system light reflected from a point object, situated at s on the lens axis, is focused onto a point in image space situated at s'. It can be seen in Fig. 2.6 that if the image plane is not positioned at s' the image formed is not that of a point object but that of a disk. This disk is the psf of an imaging system producing out-of-focus blur and the image formed, of an object situated at s, is therefore distorted or blurred in the manner described by

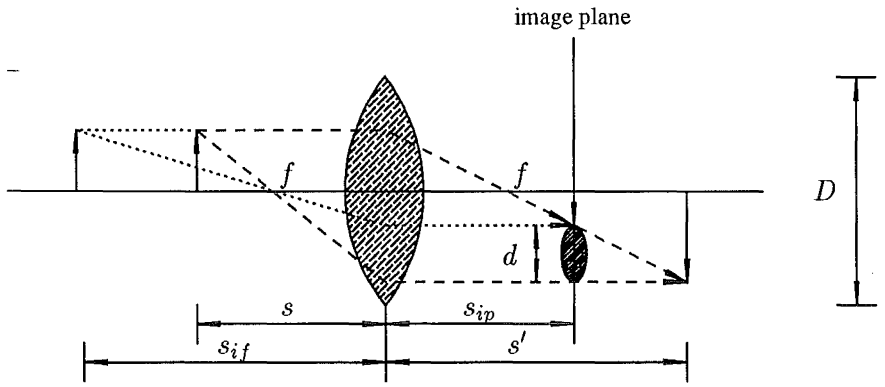

Figure 2.6 Imaging system showing how out-of-focus blur is produced by incorrectly positioning the imaging plane at s_{ip} instead of at s'. s is the position of the actual object and s_{if} is the position where the object must be positioned if it is to be correctly focused.

Eqn.(2.40). An example of the distortion that this out-of-focus blur causes is depicted in Fig. 2.7. The ideal image is shown in Fig. 2.7 (a), a psf of the out-of-focus type is shown in Fig. 2.7 (b), and the resulting blurred image image is shown in Fig. 2.7 (c).

For an imaging system where the distance, s_{ip}, between the lens and imaging plane is fixed (in practical imaging systems s_{ip} can be adjusted, over a finite range, in an attempt to bring the object into focus) the size of the psf is dependent on the distance s between the lens and the object (Biemond *et al.*, 1990). An expression for the diameter of the psf in relation to s can be determined from Fig. 2.6 by using simple geometry and the thin lens equation (Eqn.(2.30)). The geometry of Fig. 2.6 allows the diameter of the psf to be described by the expressions

$$d = \frac{D(s' - s_{ip})}{s'}, \quad f < s < s_{if}, \tag{2.41}$$

$$= \frac{D(s_{ip} - s')}{s'}, \quad s_{if} < s < \infty, \tag{2.42}$$

where d, D, s_{if}, and f are the diameter of the psf, diameter of the lens (or aperture), the distance between the object and the lens when the image is focused on the image plane (see Fig. 2.6), and the focal length of the lens respectively. Applying the thin lens equation to Fig. 2.6, the expressions

$$\frac{1}{s} + \frac{1}{s'} = \frac{1}{f}, \tag{2.43}$$

$$\frac{1}{s_{if}} + \frac{1}{s_{ip}} = \frac{1}{f}, \tag{2.44}$$

are obtained. On substituting Eqns. (2.43) and (2.44) into Eqns. (2.41) and (2.42) it is possible to eliminate the distances s' and s_{ip} from the expressions and to obtain the expressions

$$d(s) = \frac{f}{n}\left[1 - \frac{s_{if}}{s_{if} - f} + \frac{f s_{if}}{s(s_{if} - f)}\right], \quad f < s < s_{if}, \tag{2.45}$$

$$= \frac{f}{n}\left[\frac{s_{if}}{s_{if} - f} - \frac{f s_{if}}{s(s_{if} - f)} - 1\right], \quad s_{if} < s < \infty, \tag{2.46}$$

(where the f-stop $n = \frac{f}{D}$) for the diameter of the psf in terms of s and the constant s_{if} (Biemond *et al.*, 1990).

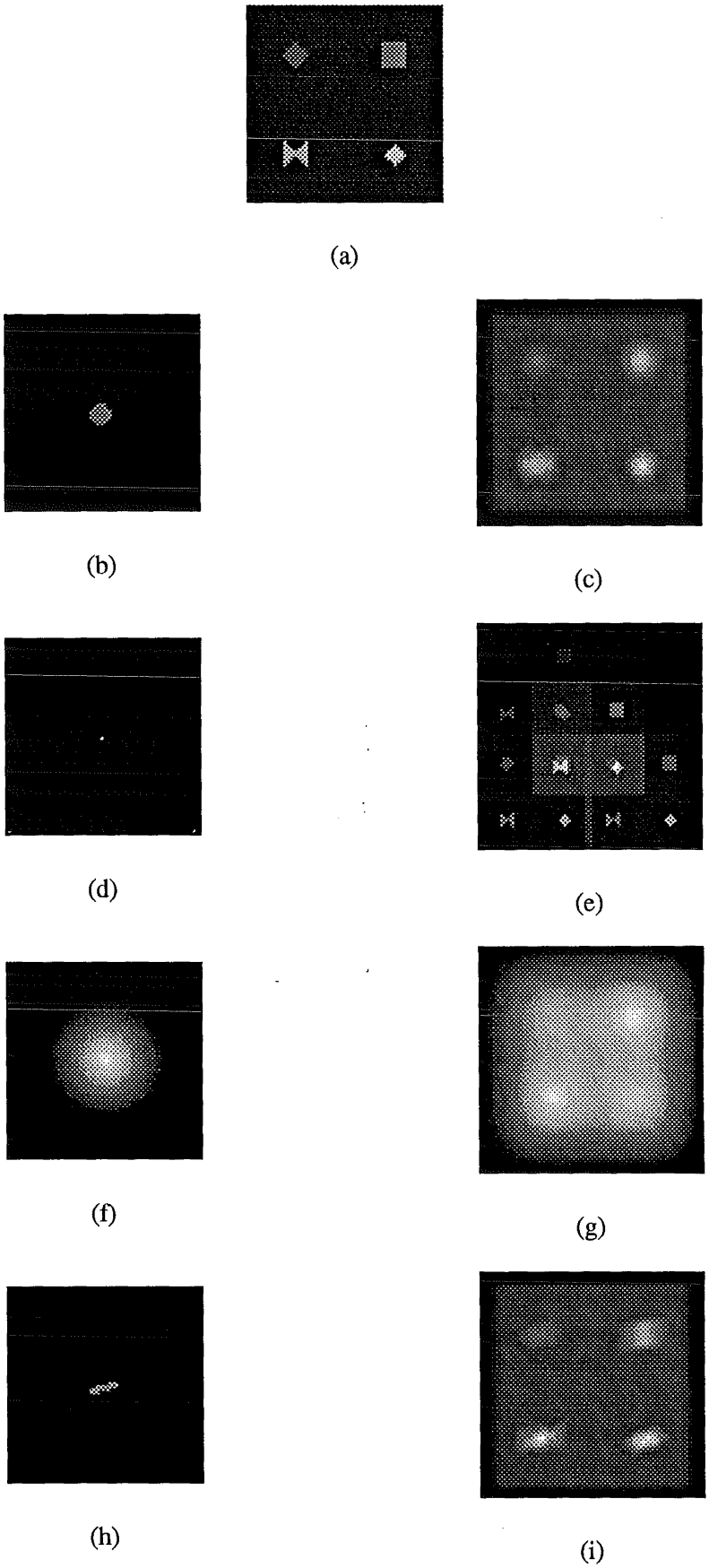

Figure 2.7 Effects of different types of blurring on the ideal image in (a), (b) out-of-focus psf, (c) out-of-focus blur with psf in (b), (d) random impulses psf, and (e) blurring due to psf in (d), (f) Gaussian psf, (g) Gaussian blurring with psf in (f), (h) motion blur psf, (i) motion blur with psf in (h). Note that the extents of the convolutions have been scaled and so the features in the convolutions (e.g. see (e)) can be smaller than in the original image (a).

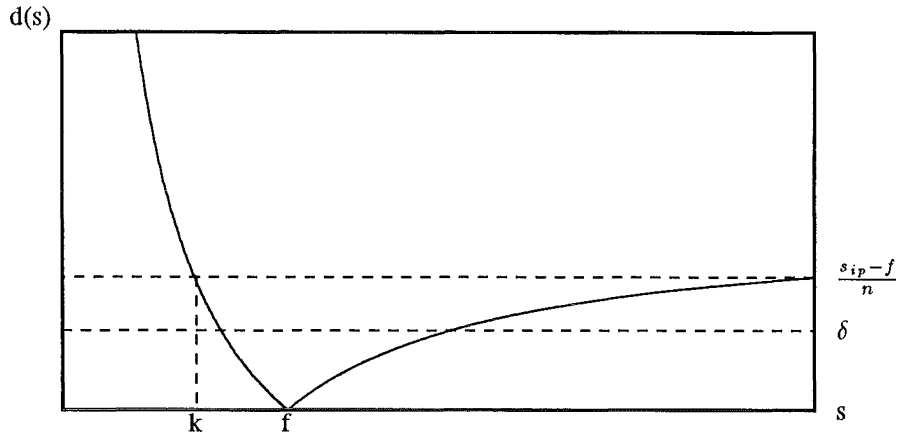

Figure 2.8 The change in the diameter of the psf as a function of the distance s between the object and the lens. δ is the resolution of the recording device, n is the f-stop, f is the focal length of the lens, and s_{ip} is the distance between the lens and the imaging plane.

The relationship between the psf diameter and object distance is shown pictorially in Fig. 2.8. This figure shows that for practical imaging systems, which use image recording devices with a finite resolution δ, there are a range of s values, known as the depth of field, over which the blur is indistinguishable, i.e. $d(s) < \delta$. It can also be seen from Fig. 2.8 that for objects that are at large distances from the lens (i.e. as $s \rightarrow \infty$) the psf has a limiting diameter of $\frac{s_{ip}-f}{n}$. In practical imaging systems it is usually possible to adjust the aperture so that $\frac{s_{ip}-f}{n} < \delta$ and the images of objects situated between $s = k$ (see Fig. 2.8) and $s = \infty$ can also be considered to be free of blurring. The effect of the blurring is, therefore, most significant when the object is situated between $s = f$ and $s = k$ where the diameter of the psf is given by Eqn.(2.45). It is necessary to note that it is assumed here that the imaging system is not diffraction limited, as either the incident light is polychromatic or the aperture is large enough to make the effects of diffraction indistinguishable, and so the resolution limit of the system due to diffraction is much less than the resolution of the recording device.

It should be noted that Eqns. (2.45) and (2.46) imply that when a 3-D object is imaged, Eqn.(2.40) cannot be used to describe the blurring since the psf's diameter now varies with the image plane position (i.e, the blurring is non-isoplanatic). In practice it is found, however, that useful results can be obtained from deblurring techniques, which invert Eqn.(2.40), by assuming that the blurring is in fact isoplanatic and including any departures from isoplanatic blurring in an additive contamination term, which is introduced later in this section (Andrews and Hunt, 1977; Bates and McDonnell, 1986). The reason useful results are obtained when this assumption is made is that it is usual, in many practical imaging situations, for very little light to be reflected from the internal structure of the object and the image is, therefore, usually formed from light which is directly reflected from points on the object's surface which can be considered to lie within a shallow volume in object space. If this object volume is now so shallow that all the points of reflection can be considered to lie on a single plane, known as the object plane, or the volume is situated at a value of s between k and infinity, the diameter of the psf does not vary significantly over the image (see Fig. 2.8). In the above two situations the assumption that the blurring is isoplanatic is therefore valid and the departures from isoplanatic behaviour included within the contamination term are very small. In a number of cases where the blurring is non-isoplanatic, it is still possible to deblur the image if there are definite isoplanatic regions in the image. In this situation the individual regions of isoplanatic

blurring, known as *isoplanatic patches*, can be separated and deblurring algorithms, using the appropriate psf for each region, can be applied to each region (Goodman, 1968; Andrews and Hunt, 1977).

The remaining two types of blurring that commonly arise in optical imaging are those caused by either random fluctuations in the refractive index of the medium (e.g. the atmosphere) or relative motion between the object and the imaging system (Andrews and Hunt, 1977; Bates and McDonnell, 1986). In the situation where an object is imaged through a turbulent medium and the exposure is small in comparison to the time interval over which the changes in the refractive index occur, the ideal image is blurred by a psf which describes the refractive index distribution at the instant of exposure. It has been found in this case that the refractive index distribution and, therefore, psf can be adequately described by a distribution of random impulses (Bates and McDonnell, 1986). The result of blurring an ideal image by a psf comprising a distribution of random impulses is shown in Fig 2.7 where the ideal image in Fig. 2.7(a) has been blurred by the psf in Fig. 2.7(d) to give the blurred image in Fig. 2.7(e). If the exposure is much larger than the time interval over which detectable changes occur in the refractive index of the medium then the psf takes the form of a Gaussian function (Andrews and Hunt, 1977; Bates and McDonnell, 1986). In Fig. 2.7(g) it is shown that a psf in the form of a Gaussian function (see Fig. 2.7(f)) behaves like a general low pass filter removing the detail from the ideal image in Fig. 2.7(a).

In the situation where the object and imaging system move relative to each other during the process of recording, the ideal image is blurred by a psf which is characterised by the relative motion (Andrews and Hunt, 1977; Bates and McDonnell, 1986; Biemond *et al.*, 1990). To enable the form of the psf describing this motion blur to be deduced it is useful to consider the image formed by the imaging system in Fig. 2.6 when the point object at s is moved whilst the recording device is being exposed. If, for the moment, it is assumed that the point object at s is moved perpendicular to the lens axis it is reasonably straightforward to show that the point object at s' is now replaced by a line at s' also perpendicular to the lens axis. On considering the linear time invariant imaging system, used previously to obtain the convolution model of image blurring, it can be seen that the psf in this case takes on the form of a line orientated in the direction of motion, since the output of the system is a line when the input is point. In general, the psf describing the motion blur is a scaled trace of the motion which induced the blur. It should be noted that if the exposure is sufficiently short or the motion is sufficiently slow then the extent of the psf could be less than the resolution of the recording device and the blurring is not detectable. The effect of motion blur on an ideal image is shown in Fig. 2.7(i) where the ideal image has been blurred by the motion psf in Fig. 2.7(h).

The second type of degradation most commonly encountered in imaging systems is distortion caused by the nonlinear response of either the image formation or recording systems. In the preceding paragraphs it was assumed that the image formation systems were LTI systems. This assumption is valid for most practical imaging systems since the departures from linearity are generally small and can be included within the additive contamination term (along with the departures from isoplanatic behaviour) to be introduced later (Bates and McDonnell, 1986). Imaging systems which use high energy X-ray absorption techniques to probe an object are an example of a group of systems where the assumption of linearity cannot be used to model the system. The reason linearity cannot be assumed is that the psf in these situations is a function of the density of the object and this precludes expressing the image formation process in the form of a linear equation (Andrews and Hunt, 1977).

The nonlinear response of most commonly used image recording devices are the most

significant contributors to the nonlinear degradation of images. Photochemical and photoelectrical sensors, such as photographic film and ccds respectively, are the main devices used to record images and both of these recording systems have a nonlinear response. For photographic images it is found that the logarithm of the density of the compound deposited on the photographic paper (i.e. silver halide) to form the image is proportional to the logarithm of the light intensity that exposed the film (Andrews and Hunt, 1977). Similarly, for photoelectric devices it is found that the logarithm of the output current of each sensor is also proportional to the logarithm of the incident light (Andrews and Hunt, 1977). Both of these devices have an output that saturates (i.e. the compound density and current has an upper limit) if the intensity of the incident light becomes too great and have a noticeable output when the incident light intensity is zero. It is customary to model the nonlinear response of the image recording system with an equation of the form

$$g'(x,y) = S\left\{b(x,y)\right\}, \tag{2.47}$$

where S is a function which models the nonlinear response of the system and $g'(x,y)$ is the recorded image (Andrews and Hunt, 1977; Biemond *et al.*, 1990). The overall nonlinear response of both of the above types of sensor can, however, be avoided to a large degree be selecting an exposure which results in an incident light intensity that corresponds to the region on the sensor response curves where the input and output intensities are approximately linearly related (i.e. the proportionality constant or gamma of the sensors is approximately equal to one). In the remainder of this thesis it is assumed that the imaging system has been configured so that the recorded image is approximately linearly related to the ideal image and any departures from linearity are so small that they can be included within the additive contamination term.

The final form of degradation that can degrade an ideal image is noise. Noise processes in the image recording system are the primary source of the noise which corrupts the ideal image. The origins and types of noise encountered are dependent on the recording process used and may consist of film-grain, thermal, photon statistical, measurement error, and quantisation noise (Andrews and Hunt, 1977; Biemond *et al.*, 1990). For a photochemical recording system the relevant type of noise is film-grain noise and this noise has its origins in the random manner in which the photochemical compound is deposited. Film-grain noise in a region of a image is statistically correlated to the intensity of the ideal image in that region, however, in most practical situations the noise process can be modelled by a Gaussian process (Andrews and Hunt, 1977). Images formed by photoelectric systems are degraded by thermal, photon statistical, and quantisation noise (Andrews and Hunt, 1977). Thermal noise has its origins in the circuits that are used to sense and process the signals from the photoelectric sensor array and arises because of the random motion that all particles above a temperature of 0 K exhibit. This noise is uncorrelated with the ideal image and can be described by a zero-mean Gaussian process. Photon statistical noise arises from the random fluctuations in the number of photons that are incident on the sensors. For high light levels this noise can be described by a Gaussian process which has a mean that is dependent on the intensity of the ideal image in the region of interest (Andrews and Hunt, 1977). For low light levels the statistics of photon statistical noise is usually approximated by a Poisson probability density function. Despite the fact that not all imaging noise is uncorrelated to the ideal image it is usual to model images corrupted by noise with the expression

$$g''(x,y) = b(x,y) + n(x,y), \tag{2.48}$$

where $n(x,y)$ is a additive Gaussian noise process which is statistically uncorrelated to the ideal image. In practice it has been found that this additive model for imaging noise is a

reasonably accurate description of the noise process and the model has been found to provide useful results (Biemond *et al.*, 1990).

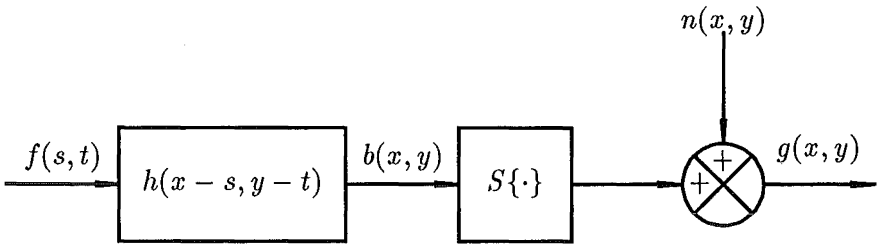

Figure 2.9 Model of the processes which degrade an ideal image $f(s,t)$. $h(x-s,y-t)$ models the blurring, $S\{\cdot\}$ models the nonlinear response of the image recording system, $n(x,y)$ models the noise, and $g(x,y)$ is the degraded image.

Using each of the above models for the various forms of image degradation it is now possible to develop an overall model of the image degradation process. Combining the models for isoplanatic blurring (Eqn.(2.40)), the nonlinear recording device response (Eqn.(2.47)), and noise (Eqn.(2.48)) the expression

$$g(x,y) = S\{f(x,y) \odot h(x,y)\} + n(x,y),$$

describing the degraded image $g(x,y)$ recorded by a nonideal imaging system, is obtained. A schematic of this image degradation model is shown in Fig. 2.9 for clarity. In the situation where there are small departures from isoplanatic blurring and the sensor response is approximately linear it is possible to model the degraded image by the expression

$$g(x,y) = f(x,y) \odot h(x,y) + c(x,y), \tag{2.49}$$

where the term $c(x,y)$ is an additive contamination term which now describes the departures from isoplanatic blurring and linearity in addition to the noise (Andrews and Hunt, 1977; Bates and McDonnell, 1986; Biemond *et al.*, 1990). Throughout the remainder of this thesis it is assumed that the degradation on the images to be considered can be adequately described by Eqn.(2.49) and the algorithm in §5.1 has therefore been developed with this model in mind.

2.4 CONVOLUTION AND CORRELATION

The similar processes of convolution and correlation have been found to be extremely valuable in engineering science, where they have been used extensively to facilitate the development of mathematical models of physical systems and to enable the distortion in these systems to be detected, modelled, and eliminated (or at least reduced). In this section the mathematical formulae which describe both the convolution and correlation processes are introduced in turn and the concepts behind the formulae and their application to system modelling is discussed.

Of the two processes, the process which primarily forms the basis for many mathematical models of a system is convolution. The convolution operation can allow a *black box* model of a system to be developed that describes the output of the system in terms of its input and a function known as the system transfer function in Fourier space and the impulse response in the spatial domain. Control systems (Sinha and Kuszta, 1983), imaging systems (see §2.3.2) (Bates and McDonnell, 1986; Biemond *et al.*, 1990), communication systems (Haykin, 1983), and the human speech process (Witten, 1982) are a few examples of areas where convolution models

have provided useful models of systems. To be able to use the convolution operator to model a system it is necessary for the system to be, or approximate, a linear time invariant (LTI) system. A LTI system is a system in which superposition applies and the parameters which characterise the system do not change with time or the form of the input (Goodman, 1968; Andrews and Hunt, 1977; Brigham, 1988). The superposition property, which is only displayed by linear systems, is the requirement that a system's response to an input comprising a summation of a number of separate signals is equal to the summation of the responses that would be produced by the signals individually. Expressing this property mathematically; all LTI systems obey the relationship

$$y(\mathbf{x}) = \mathcal{L}\{k_1 q(\mathbf{x}) + k_2 r(\mathbf{x})\} = k_1 \mathcal{L}\{q(\mathbf{x})\} + k_2 \mathcal{L}\{r(\mathbf{x})\}$$

where $\mathcal{L}\{\cdot\}$ is the function describing the system response, $q(\mathbf{x})$ and $r(\mathbf{x})$ are the separate complex input signals, $y(\mathbf{x})$ is the output, and k_1 and k_2 are complex constants.

Before introducing the convolution operator it is first necessary to introduce the Dirac delta or unit impulse function. The Dirac delta function $\delta(\mathbf{x})$, is a function for which (Bracewell, 1978)

$$\begin{aligned} \delta(\mathbf{x}-\boldsymbol{\tau}) &= 0 \qquad \mathbf{x} \neq \boldsymbol{\tau} \\ (K)\int_{-\infty}^{\infty} \delta(\mathbf{x}-\boldsymbol{\tau})d\mathbf{x} &= 1, \end{aligned}$$

where $\boldsymbol{\tau}$ is the position vector of the delta function in space. A useful property of the delta function is the sifting property

$$(K)\int_{-\infty}^{\infty} r(\mathbf{x})\delta(\mathbf{x}-\boldsymbol{\tau})d\mathbf{x} = r(\boldsymbol{\tau}), \tag{2.50}$$

which allows a single value of a function, in this case $r(\mathbf{x})$ at $\mathbf{x} = \boldsymbol{\tau}$, to be sifted out from all the possible values of the function (Bracewell, 1978). By changing the roles of $\mathbf{x}$ and $\boldsymbol{\tau}$ in Eqn.(2.50), the equation emphasises that any function $r(\mathbf{x})$ can be considered as the superposition of impulses at positions $\boldsymbol{\tau}$, weighted according to $r(\boldsymbol{\tau})$.

As demonstrated in §2.3.2, for the case of an imaging system, the principle that any signal can be represented as a summation of weighted impulses and the principle of superposition, can be used to obtain an expression for the output, $y(\mathbf{x})$, of a LTI system (and spatially invariant system in the case of an imaging system) in terms of its input, $r(\mathbf{x})$, and its impulse response h(**x**). Thus, the output of any LTI system can be described by an expression of the form

$$\begin{aligned} y(\mathbf{x}) &= (K)\int_{-\infty}^{\infty} r(\boldsymbol{\tau})h(\mathbf{x}-\boldsymbol{\tau})d\boldsymbol{\tau} \\ &= r(\mathbf{x}) \odot h(\mathbf{x}), \end{aligned} \tag{2.51}$$

which is known as the convolution integral and where $\odot$ denotes the convolution of the two functions. One of the desirable properties of the convolution integral is that on Fourier transforming both sides of Eqn.(2.51) the convolution theorem

$$Y(\mathbf{u}) = R(\mathbf{u})H(\mathbf{u}), \tag{2.52}$$

where $H(\mathbf{u})$ is the system transfer function alluded to earlier, is obtained (see §2.6) (Bates and McDonnell, 1986). It can be seen from Eqn.(2.52) that convolution in Fourier space can be performed with a simple multiplication. This property is the basis for many deconvolution algorithms (see §3.2).

Related to convolution is the process of correlation which is primarily used to determine whether two signals are related to each other mathematically (i.e. to evaluate the independence of two signals). If the correlation function indicates that two signals are related to each other, it also then provides some information as to the nature of the relationship that connects the signals. The correlation of two individual signals, known as the cross-correlation function, is defined by the expression

$$\begin{aligned} z_{rh}(\mathbf{x}) &= \int_{-\infty}^{\infty} r^*(\tau)h(\mathbf{x}+\tau)d\tau \\ &= r(\mathbf{x}) \star h(\mathbf{x}), \end{aligned} \tag{2.53}$$

where $\star$ denotes the process of correlation (Goodman, 1968). On applying the Fourier transform to Eqn.(2.53) the cross-correlation theorem

$$Z_{RH}(\mathbf{u}) = R^*(\mathbf{u}^*)H(\mathbf{u}), \tag{2.54}$$

which is derived in §2.6, is obtained. The cross-correlation theorem shows that the cross-correlation of two signals principally reduces to a series of multiplications in Fourier space, consequently once the multiplications have been performed, the cross spectrum Z_{RH} can be simply inverse Fourier transformed to obtain the cross-correlation function. The expression in Eqn.(2.53) shows that the correlation function is basically a function describing the area under the product of a signal and a time shifted version of another signal, as a function of the time shift $\mathbf{x}$. If two signals are completely independent then $z_{rh} = 0 \ \forall \ \mathbf{x}$, where as, if two signals are related z_{rh} has a maximum at the value of $\mathbf{x}$ where the signals are most similar. Another important attribute of cross-correlation, that can be used to find models of linear systems, is the property that the cross-correlation of a system's input and output provides the impulse response of the system if the input to the system is white noise (Lange, 1967). The above characteristics of the cross-correlation process make it a valuable tool for detecting and measuring any distortion introduced into a system and for developing models of an overall system or distortion process. For a more detailed discussion on the applications and properties of cross-correlation the reader is referred to Lange (1967) and Bendat and Pierslo (1980).

An important class of correlation function is produced when a signal is correlated with itself. This correlation function is known as the auto-correlation function and is defined by the expression

$$\begin{aligned} z_{rr}(\mathbf{x}) &= \int_{-\infty}^{\infty} r^*(\tau)r(\mathbf{x}+\tau)d\tau \\ &= r(\mathbf{x}) \star r(\mathbf{x}). \end{aligned} \tag{2.55}$$

The similarity between convolution and correlation is most evident with the auto-correlation as it is possible to rewrite Eqn.(2.55) in the form of a convolution

$$z_{rr}(\mathbf{x}) = r^*(\mathbf{x}) \odot r(-\mathbf{x}). \tag{2.56}$$

Using the same derivation in §2.6 that is used to derive the cross spectrum, it is also possible to transform Eqn.(2.55) into Fourier space giving

$$Z_{RR}(\mathbf{u}) = R^*(\mathbf{u}^*)R(\mathbf{u}), \tag{2.57}$$

which is known as the auto-correlation theorem. On the real plane in Fourier space $\mathbf{u}^* = \mathbf{u}$ and Eqn.(2.57) reduces to

$$\begin{aligned} Z_{RR}(\mathbf{u}) &= R^*(\mathbf{u})R(\mathbf{u}), \\ &= |R(\mathbf{u})|^2, \end{aligned} \tag{2.58}$$

which shows that the spectrum of a signal's auto-correlation, for real $\mathbf{u}$, is equal to the signal's power spectrum. This property has important ramifications in phase retrieval and is the basis for the new algorithm developed in §6.1.2.

2.5 FOURIER TRANSFORMS

One of the most useful mathematical tools available to scientists and engineers is the Fourier transform which maps a K-dimensional signal between the K-dimensional spatial domain (i.e. image space and/or time) and the K-dimensional frequency domain (i.e. Fourier space). The usefulness of the Fourier transform primarily stems from a number of properties of the Fourier transform, introduced in §2.6, which relate the form of a function in one domain to that of its transform. In many instances, because of the properties of the Fourier transform, it has been found that Fourier transforming a mathematical description of a problem may result in a more obvious or significantly simplified solution. Two excellent examples of problems which can be greatly simplified by Fourier transforming their mathematical descriptions are the deblurring and phase retrieval problems which are formally introduced in §3.1 and whose frequency domain solutions were alluded to in §2.4. In addition, to the useful characteristics of the Fourier transform described above, the Fourier transform also has the valuable property that the behaviour of a number of systems, such as lenses and antennae, can be modelled by using the Fourier transform relationship (see §2.3.1.2).

2.5.1 Continuous Fourier Transforms

The Fourier transform of the continuous, and possibly complex, K-dimensional function a(**x**) is defined by the expression

$$\mathsf{F}\{a(\mathbf{x})\} = A(\mathbf{u}) = (K)\int_{-\infty}^{\infty} a(\mathbf{x})\exp^{i2\pi\mathbf{u}\cdot\mathbf{x}}d\mathbf{x} \tag{2.59}$$

where $\mathsf{F}\{\cdot\}$ denotes the Fourier transform operation and $A(\mathbf{u})$ is known as the spectrum of $a(\mathbf{x})$. The inverse Fourier transform, defined by

$$\mathsf{F}^{-1}\{A(\mathbf{u})\} = a(\mathbf{x}) = (K)\int_{-\infty}^{\infty} A(\mathbf{u})\exp^{-i2\pi\mathbf{u}\cdot\mathbf{x}}d\mathbf{u}, \tag{2.60}$$

where $\mathsf{F}^{-1}\{\cdot\}$ denotes the inverse Fourier transform, allows the function to be transformed back into the spatial domain. If both the integrals in Eqns. (2.59) and (2.60) converge, which they do for most functions encountered in practical situations, the equations define a unique invertible mapping between the spatial and frequency domains (Bracewell, 1978; Brigham, 1988). In this thesis the invertible mapping between the two domains is denoted by the expression

$$a(\mathbf{x}) \overset{FT}{\Longleftrightarrow} A(\mathbf{u}),$$

where the above functions $a(\mathbf{x})$ and $A(\mathbf{u})$ are known as a Fourier transform pair. It should be noted that some authors define the forward and inverse Fourier transform pair with and without the minus sign respectively, however provided that consistency is maintained this is of no consequence.

2.5.2 Fourier Transforms of Discrete Images

In most real world situations the quantities that are of interest to physical scientists are continuous quantities. Depending on the nature of the quantity of interest, and the purpose for which it is required, it is often necessary, or only feasible, to measure the quantity at a finite number of points in time or space. The process of measuring the quantity at a finite number of points in space to obtain a discrete representation of the quantity is known as sampling. One of the main reasons real world quantities must be sampled is that digital computers, which can only operate on a finite set of numbers, are often used to further process these quantities. The sampling method which must be adopted to ensure that the set of samples accurately represent the quantity is discussed in §2.5.4. This section is primarily concerned with deriving an expression for the continuous Fourier transform of a sampled image. This expression is derived here as it is required later in §2.7 and Chapter 4 where the relationship between the Fourier transform and Z transform and the concept of the zero sheet are introduced respectively.

In most practical imaging situations, because of the finite resolution of photographic film, the finite size of the ccd sensor array, or the need to process the image with a digital computer, the image obtained is a sampled version of the ideal image. In addition, before entering photographic images into a computer they must be resampled using a scanning device which converts the image into a form that can be directly entered into the computer and reduces the data set to a more manageable size. It is assumed here, as in §2.3.2, that the resolution limit of the imaging system due to diffraction is much less than the distance between samples. These sampled images, in the 2-D case, can be represented by an N by M matrix whose elements (or pixels) $f[n,m]$ are a representation of the ideal image. In the situation where these pixels are of equal size in both spatial dimensions the sampled image $f_s(x,y)$ can be represented by the expression

$$f_s(x,y) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f[n,m]\delta(x-n\epsilon)\delta(y-m\epsilon) \tag{2.61}$$

where ϵ is the pixel spacing and the subscript s denotes that the image is sampled. In the following chapters, which primarily deal with pixellated images, the subscript s is dropped. To obtain the Fourier transform of this sampled image, Eqn.(2.61) is substituted into the Fourier transform relationship (Eqn.(2.59)) giving

$$F(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f[n,m]\delta(x-n\epsilon)\delta(y-m\epsilon)\exp^{i2\pi(ux+vy)}dxdy.$$

Changing the order of integration and summation in the above equation and expanding the resulting expression gives (Bracewell, 1978)

$$\begin{aligned}
F(u,v) \;=\; & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f[0,0]\delta(x)\delta(y)\exp^{i2\pi(ux+vy)}dxdy + \\
& \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f[0,1]\delta(x)\delta(y-\epsilon)\exp^{i2\pi(ux+vy)}dxdy + \\
& \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f[1,0]\delta(x-\epsilon)\delta(y)\exp^{i2\pi(ux+vy)}dxdy + \\
& \ddots + \\
& \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f[N-1,M-1]\delta(x-(N-1)\epsilon)\delta(y-(M-1)\epsilon)\exp^{i2\pi(ux+vy)}dxdy.
\end{aligned}$$

Evaluating the integral in the above series and reducing the series gives

$$\begin{aligned} F(u,v) &= \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f[n,m]\exp^{i2\pi(un\epsilon+vm\epsilon)} \\ &= \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f[n,m]\exp^{i2\pi un\epsilon}\exp^{i2\pi vm\epsilon}, \end{aligned} \tag{2.62}$$

which is the continuous Fourier transform of a sampled image.

2.5.3 Discrete Fourier Transforms

To compute the Fourier transform on a digital computer the spatial domain signal and its spectrum must both be sampled and be of finite extent. If the expression for the continuous Fourier transform is modified so that it has all of the above properties, an expression known as the discrete Fourier transform (DFT), which can be used to transform sampled signals into the sampled frequency domain, is obtained. A full development of the DFT is beyond the scope of this thesis and for this reason only the basic outline of the steps required to derive the DFT, along with the graphical development in Fig. 2.10, are provided since this enables some of the relevant pitfalls of the DFT to be highlighted. A detailed explanation of the derivation of the DFT from the continuous Fourier transform can be found in many Fourier transform texts such as Brigham (1988).

Following Brigham, the DFT of the continuous signal $a(x)$ in Fig. 2.10 (a), which for reasons of clarity is assumed to be a 1-D signal, can be determined as follows. The first step is to convert $a(x)$ into the discrete signal $a_s(x)$ by sampling it with a series of equally spaced delta functions (see Fig. 2.10 (b)). If the spacing between the individual delta functions is T_s then the sampled signal can be described by the expression

$$\begin{aligned} a_s(x) &= a(x)\sum_{n=-\infty}^{\infty}\delta(x-nT_s), \\ &= \sum_{n=-\infty}^{\infty} a(nT_s)\delta(x-nT_s). \end{aligned}$$

In the frequency domain the effect of sampling can be determined by considering the convolution theorem introduced in §2.4 and the duality property of the Fourier transform which is inherent in Fourier transform pair relationship introduced in §2.5.1. These properties of the Fourier transform indicate that multiplication in the time domain is equivalent to convolution in the frequency domain and hence, the spectrum in Fig. 2.10 (c) of the sampled signal is obtained by convolving the spectra of the series of delta functions with that of the continuous signal. It should be noted that the spectrum of the sampled signal is a periodic function which has a period equal to $\frac{1}{T_s}$ (i.e. the inverse of the spacing between the spatial domain delta functions). As described earlier a digital computer can only operate on a finite set of numbers, hence, it is now necessary to truncate the series representing the signal to an appropriate size of N samples. To truncate $a_s(x)$ it is multiplied by what is known as a windowing function, which in this instance is assumed to be the rectangular function Rect(x) in Fig. 2.10 (d), giving the truncated signal

$$a_s(x)\text{Rect}(x) = \sum_{n=0}^{N-1} a(nT_s)\delta(x-nT_s).$$

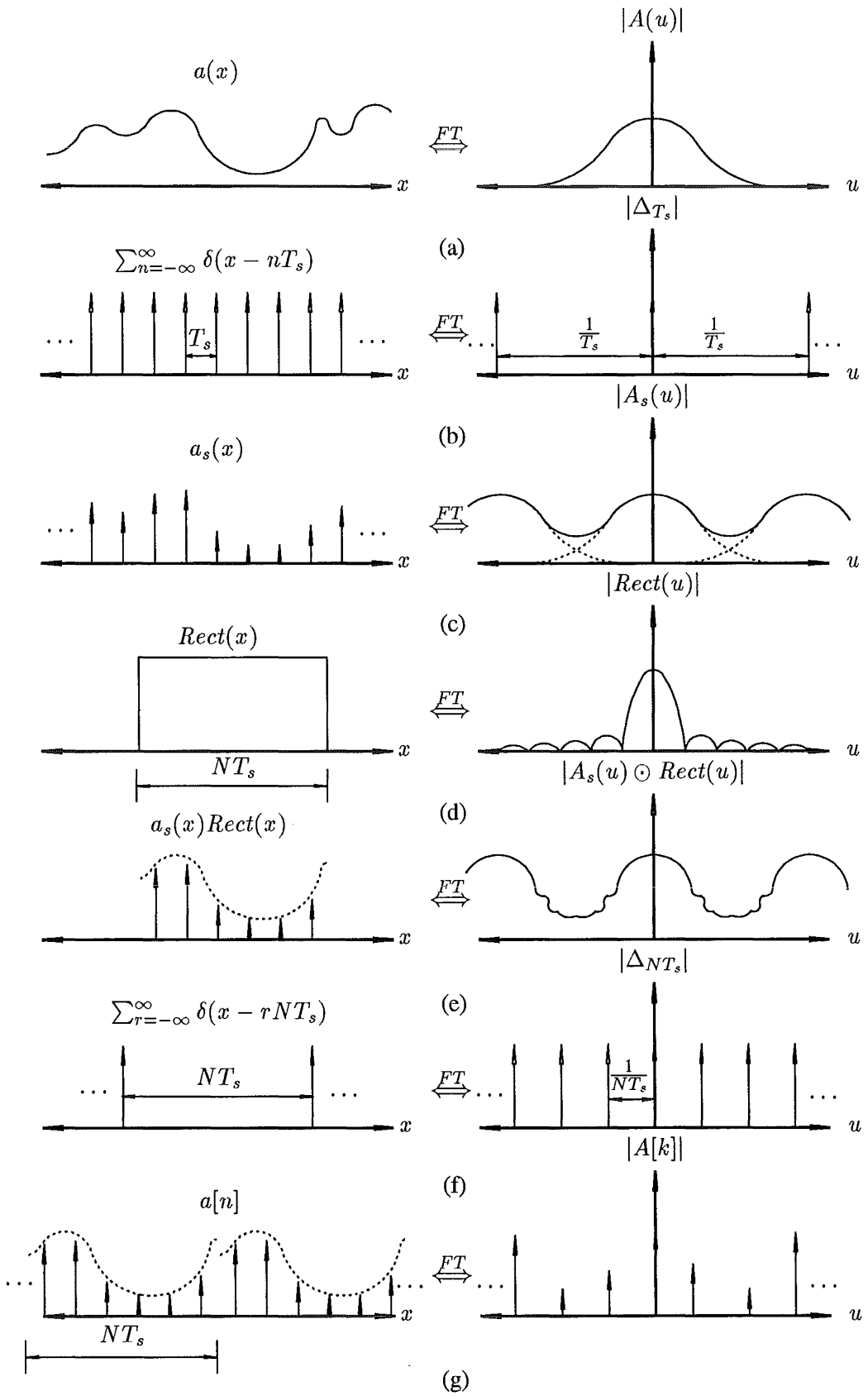

Figure 2.10 Graphical development of the Discrete Fourier Transform. (a)The continuous signal to be discrete Fourier transformed, (b) series of delta functions used to sample the signal in (a) in the spatial domain, (c) a sampled version of the signal in (a), (d) truncation function used to truncate the signal in (c), (e) the truncated version of the signal in (c), (f) series of delta functions used to sample the spectrum of the truncated function in (e), and (f) the DFT of the signal in (a). Note that the DFT of the signal is periodic in both domains.

Truncating the signal in the time domain has the effect of distorting the spectrum of $a_s(x)$, as is shown in Fig. 2.10 (e), since the spectrum of the truncated and sampled signal is the convolution of the spectrum of the truncation function and $A_s(u)$.

The next step required to obtain the DFT is to sample the spectrum of the truncated and sampled signal. This is again done by multiplying the function to be sampled by a series of equally spaced delta functions denoted $\Delta_{NT_s}(u)$. To ensure that the resulting convolution in the time domain does not result in any distortion of the time domain function these impulses are separated in the frequency domain by the distance $\frac{1}{NT_s}$ (see Fig. 2.10 (e)). In the time domain the resulting signal is described by the expression

$$\begin{aligned}\{a_s(x)\text{Rect}(x)\} \odot \delta_{NT_s}(x) &= \left\{\sum_{n=0}^{N-1} a(nT_s)\delta(x-nT_s)\right\} \odot \sum_{r=-\infty}^{\infty} \delta(x-r(NT_s)) \\ &= \sum_{r=-\infty}^{\infty}\sum_{n=0}^{N-1} a(nT_s)\delta(x-nT_s-r(NT_s)), \qquad (2.63)\end{aligned}$$

which is a periodic function with a period NT_s. To obtain the DFT the knowledge that the Fourier transform of a periodic function $s_p(x)$, with a period T_0, is given by the expression

$$S_p\left[\frac{k}{T_0}\right] = \sum_{k=-\infty}^{\infty} \alpha[k]\delta\left(u-\frac{k}{T_0}\right),$$

where (Brigham, 1988)

$$\alpha[k] = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} s_p(x)\text{exp}^{i2\pi kx/T_0}dx \qquad k = 0, \pm 1, \pm 2, \ldots,$$

is used to obtain the Fourier transform of the function in Eqn.(2.63) (see Fig. 2.10 (g)). On performing this Fourier transform the DFT relationship

$$A[k] = \sum_{n=0}^{N-1} a[n]\text{exp}^{i2\pi kn/N} \qquad k = 0, 1, \ldots, N-1, \qquad (2.64)$$

where $a[n]$ and $A[k]$ are the samples of the spatial and Fourier domain functions respectively, is obtained (Bracewell, 1978; Brigham, 1988). The inverse DFT (IDFT) is defined by the relationship

$$a[n] = \sum_{k=0}^{N-1} A[k]\text{exp}^{-i2\pi kn/N} \qquad n = 0, 1, \ldots, N-1. \qquad (2.65)$$

On substituting Eqn.(2.64) into Eqn.(2.65) it is found that both sides of the resulting equation are indeed equal and this demonstrates that Eqns. (2.64) and (2.65), as with the continuous Fourier transform, define a unique invertible relationship (Bracewell, 1978). In 2-D the DFT and IDFT are defined by the relationships

$$\begin{aligned}A[k,l] = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} a[n,m]\text{exp}^{i2\pi kn/N}\text{exp}^{i2\pi lm/M} \qquad & k = 0, 1, \ldots, N-1 \\ & l = 0, 1, \ldots, M-1, \qquad (2.66)\end{aligned}$$

and

$$\begin{aligned}a[n,m] = \sum_{l=0}^{M-1}\sum_{k=0}^{N-1} A_{k,l}\text{exp}^{-i2\pi kn/N}\text{exp}^{-i2\pi lm/M} \qquad & n = 0, 1, \ldots, N-1 \\ & m = 0, 1, \ldots, M-1, \qquad (2.67)\end{aligned}$$

respectively (Brigham, 1988). To efficiently compute the DFT and IDFT, the fast Fourier transform, which makes use of redundancies in Eqns. (2.64) and (2.65), is used (Brigham, 1988).

It is emphasised at this point that the spectrum obtained by DFT is not simply a sampled version of the spectrum that would be obtained if the continuous Fourier was used, but an approximation to the spectrum that would be obtained with the continuous transform (Brigham, 1988). The process of sampling and truncating the spatial domain function distorts the spectrum of the function in a manner that is dependent on the bandwidth of the signal that the function represents (for a set spatial domain sampling interval) and the form of the window function used to truncate the data set. The distortion which is introduced into the spectrum because of the form of the window is known as spectral leakage and this is now described. The distortion that is bandwidth and sampling interval dependent, which is known as aliasing, is described in the following section. It was shown earlier that truncation of the sampled function results in the spectrum of the function being distorted since the spectrum of the sampled function has been convolved with that of the window function. To determine what effect this distortion has on the function's spectrum assume for the moment that the function's spectrum consists of a single delta function situated at $u = u_\delta$. On truncating the spatial domain function the delta function is replaced by the windowing functions spectrum, centred about u_δ, as a result of the convolution. The spectra of all the usual windowing functions have a form similar to that of the rectangular window depicted in Fig. 2.10 (d), which consists of a main lobe containing the majority of the energy and a number of sidelobes which diminish in size as $|u|$ increases (Brigham, 1988). Truncating the time domain function, consequently, results in a spectrum which has more significant frequency components at higher frequencies than before (due to the side lobes) and due to the width of the main lobe also has the majority of its energy *smeared* over a larger region of Fourier space than before. With the rectangular window the broadening of the significant features of the function's spectrum does not distort the spectrum as much as the appearance of the significant frequency components at higher frequencies. In order to reduce the significance of these higher frequency components, which are an artifact of the DFT, a different windowing function, which has less significant sidelobes, must be used. It is found, however, that as the sidelobes become less significant the width of the main lobe broadens reducing the frequency resolution of the DFT (Brigham, 1988). A trade off between the significance of the high frequency artifacts and resolution must therefore be made in practice (Brigham, 1988).

2.5.4 Sampling Images

In the situation where it is necessary to sample a continuous function, or resample a function that has already been sampled, the interval between the delta functions comprising the sampling function must be chosen appropriately if the sampled function is to be a accurate representation of the continuous function. The effect of choosing an inappropriate sampling interval can be determined graphically by considering the effect of the sampling process on the spectrum of the continuous function to be sampled. In §2.5.3 it was demonstrated that sampling a continuous function, by multiplying the function with a series of delta functions, has the effect of convolving the spectrum of the function being sampled with the spectrum of a series of delta functions. The spectrum of the sampled function is therefore a periodic function consisting, of the continuous signal's spectrum superimposed on each of the delta functions in the frequency domain which represent the spectrum of the sampling function (see Fig. 2.10(c)). If the image or signal, which the continuous function represents, has a bandwidth BW then the sampling interval T_s must be less than $\frac{1}{2BW}$ (i.e. the Nyquist interval), otherwise the individual continuous signal

spectra, which form the sampled signals spectrum, overlap in Fourier space as shown in Fig. 2.10 (c). The overlap in Fourier space superimposes high frequency components on to low frequency components, thereby, distorting the resulting spectrum so that one period of the sampled signals spectrum is no longer equivalent to the continuous signal's spectrum. This frequency domain distortion, which has the effect that a sampled high frequency signal appears to be a sampled lower frequency signal, is known as aliasing. Provided that $T_s < \frac{1}{2BW}$, or equivalently if the overlap in Fourier space occurs below the level of the inevitable noise, the effect of the aliasing is negligible and the sampled signal is an accurate representation of the continuous signal. In the case where the function being sampled is an image it is, of course, necessary to ensure the sampling is sufficiently fine in each spatial direction so that no aliasing occurs in the corresponding frequency domain directions.

In the frequency domain there is an requirement, equivalent to that in the spatial domain, which defines the size of the sampling interval that must be used in Fourier space if an image is to be reconstructed from its Fourier transform. This relationship was alluded to in §2.5.3 where the sampling interval in the frequency domain for the DFT was made equal to $1/NT_S$, where NT_S was the extent of the truncated function in the spatial domain, so that the resulting convolution did not result in an overlap of the truncated spatial domain functions. In the situation where the spatial domain function is a compact image, which has extents L_x and L_y in the x and y directions respectively, the Fourier transform of the image must be sampled at intervals of $1/L_x$ and $1/L_y$ in the corresponding Fourier space coordinates if the samples are going to be used to obtain an undistorted representation of the image. This result can be obtained from the same reasoning used to obtain the spatial domain sampling interval, except that the role of image space and Fourier space are now reversed and BW is now equivalent to $\frac{L_{X(Y)}}{2}$.

It is possible to derive the above frequency domain sampling requirement mathematically by using the method of Bates and McDonnell (1986) where the compact image $a(x, y)$, with extents L_x and L_y, is described by a Fourier series. On substituting the Fourier series into the continuous Fourier transform and evaluating the resulting expression it is found that

$$A\left(\frac{s}{L_x}, \frac{t}{L_y}\right) = A_{s,t},$$

where A, $A_{s,t}$, and s and t, are the Fourier transform of the image, the Fourier coefficients and the indices of the Fourier coefficients respectively (Bates and McDonnell, 1986). The above relationship confirms that an image is completely defined by its sampled Fourier transform if the sampling intervals in Fourier space are $1/L_x$ and $1/L_y$.

2.6 FOURIER TRANSFORM PROPERTIES

In this section a number of properties of the Fourier transform are introduced. The list of Fourier transform properties in Table 2.1 relate either a spatial domain operator to the equivalent operator in the frequency domain or relate the form of a spatial domain function to the form of its Fourier transform. The Fourier transform properties are introduced here as they are required throughout this thesis to enable other relationships to be developed or important concepts to be introduced. Two properties of the Fourier transform listed in Table 2.1, the convolution and correlation theorems, are used extensively throughout this thesis and were first introduced in §2.4. Derivations of the properties in Table 2.1 can be found in a number of general texts dealing with either Fourier transforms, communications or signal processing

(Goodman, 1968; Bracewell, 1978; Haykin, 1983; Brigham, 1988). Most texts only consider real spatial domain functions (i.e. $r(\mathbf{x}) = r^*(\mathbf{x})$) and the real plane in Fourier space (i.e. $\mathbf{u} = \mathbf{u}^*$), where as the properties listed in Table 2.1 have been generalised so that they apply for both complex spatial domain functions and complex $\mathbf{u}$. It should also be noted that the definition used here for the Fourier transform differs from that used in most texts (see §2.5.1). For this reason the sign of the exponential in the shift and modulation properties and the sign of the complex factor in the differentiation and integration properties in Table 2.1 is the opposite to that found in most texts.

FOURIER TRANSFORM PROPERTIES

Property	Spatial Domain		FrequencyDomain
Linearity	$k_1 r(\mathbf{x}) + k_2 q(\mathbf{x})$	$\overset{FT}{\Longleftrightarrow}$	$k_1 R(\mathbf{u}) + k_2 Q(\mathbf{u})$
Duality	$r(\mathbf{x})$	$\overset{FT}{\Longleftrightarrow}$	$R(-\mathbf{u})$
Scaling	$r(k_1 \mathbf{x})$	$\overset{FT}{\Longleftrightarrow}$	$\frac{1}{k_1} R(\frac{\mathbf{u}}{k_1})$
Shifting	$r(\mathbf{x} - \tau)$	$\overset{FT}{\Longleftrightarrow}$	$\exp^{i2\pi \mathbf{u}\cdot\tau} R(\mathbf{u})$
Modulation	$\exp^{-i2\pi \mathbf{x}\cdot\tau} r(\mathbf{x})$	$\overset{FT}{\Longleftrightarrow}$	$R(\mathbf{u} - \tau)$
Convolution	$r(\mathbf{x}) \odot q(\mathbf{x})$	$\overset{FT}{\Longleftrightarrow}$	$R(\mathbf{u})Q(\mathbf{u})$
Cross-correlation	$r(\mathbf{x}) \star q(\mathbf{x})$	$\overset{FT}{\Longleftrightarrow}$	$R^*(\mathbf{u}^*)Q(\mathbf{u})$
Conservation of Energy (Parseval's or Rayleigh's Theorem)	$\int \lvert r(\mathbf{x})\rvert^2 \, d\mathbf{x}$	$\overset{FT}{\Longleftrightarrow}$	$\int \lvert R(\mathbf{u})\rvert^2 \, d\mathbf{u}$
Differentiation	$\frac{d}{d\mathbf{x}} r(\mathbf{x})$	$\overset{FT}{\Longleftrightarrow}$	$-i2\pi \mathbf{u} R(\mathbf{u})$
Integration	$\int_{-\infty}^{t} r(\mathbf{x}) d\mathbf{x}$	$\overset{FT}{\Longleftrightarrow}$	$\frac{-1}{i2\pi \mathbf{u}} R(\mathbf{u})$
Conjugate Functions	$r^*(\mathbf{x})$	$\overset{FT}{\Longleftrightarrow}$	$R^*(-\mathbf{u}^*)$
Real in Spatial Domain	$\Im\{r(\mathbf{x})\} = 0$	$\overset{FT}{\Longleftrightarrow}$	$R(\mathbf{u}) = R^*(-\mathbf{u}^*)$

Table 2.1

The convolution and correlation theorems are not usually derived in general texts for the case where both the spatial domain functions and frequency domain variables are complex. The importance of these theorems to the work in this thesis and the lack of a suitable derivation elsewhere suggests it would be beneficial if the basic outline for deriving the cross-correlation theorem was provided here. Once the cross-correlation theorem has been derived the similarity between the processes of convolution and correlation makes the derivation of the convolution theorem completely mechanical and for this reason only the derivation of the cross-correlation theorem is presented here. Following the method of Bates and McDonnell (1986) the cross-correlation theorem is derived by expressing the 1-D Fourier transform of the cross-correlation in the form

$$\mathsf{F}\{r(x) \star q(x)\} = \int \int_{-\infty}^{\infty} r^*(\tau) q(x+\tau) \exp^{i2\pi u x} d\tau dx. \tag{2.68}$$

Using the inverse Fourier transform relationship (Eqn.(2.60)), $r^*(\tau)$ and $q(x+\tau)$ are expressed in terms of their respective spectra and the dummy Fourier space variables α and β are used in each of the respective expressions. Substituting these expressions for $r^*(\tau)$ and $q(x+\tau)$ into

Eqn.(2.68) gives

$$\mathsf{F}\{r(x) \star q(x)\} = \int\int\int\int_{-\infty}^{\infty} R^*(-\alpha^*)Q(\beta)\exp^{i2\pi(-\alpha\tau-\beta x-\beta\tau+ux)}d\alpha d\beta d\tau dx,$$

where $r^*(x) \overset{FT}{\Longleftrightarrow} R^*(-u^*)$ as given in Table 2.1. On integrating the above expression first with respect to x and then τ the above expression reduces to

$$\mathsf{F}\{r(x) \star q(x)\} = \int\int_{-\infty}^{\infty} R^*(-\alpha^*)Q(\beta)\delta(u-\beta)\delta(-\alpha-\beta)d\alpha d\beta.$$

The cross-correlation theorem,

$$r(x) \star q(x) \overset{FT}{\Longleftrightarrow} R^*(u^*)Q(u),$$

is then obtained after evaluating the remaining integrals with respect to α and β. Generalisation of the cross-correlation theorem to K-dimensions is straightforward.

2.7 Z TRANSFORMS AND THEIR PROPERTIES

The Z transform is a transformation that is used to map a discrete function, in the time or spatial domains, into another domain known as Z space, where the transformed function is now a continuous function of the Z space coordinates. The Z transform is an invaluable tool that can be used to analyse and model discrete functions and discrete LTI systems. The manner in which the Z transform is used to analyse and model discrete functions and systems is very similar to the manner in which the continuous Fourier transform is used to model continuous functions and continuous LTI systems. As with the Fourier transform in the continuous situation, transformation of a discrete problem into the Z transform equivalent of the frequency domain (i.e. Z space) can make the acquisition of a solution to a problem easier or can greatly simplify the computation required to obtain a solution.

In this thesis the Z transform is used to transform discrete images from image space into Z space. As the images dealt with here are compact, the Z transform of the discrete $N \times M$ pixel image, with pixel values $a[n,m]$ (see §2.5.2), is defined by the 2-D, one sided, finite length, Z transform

$$\boldsymbol{Z}\{a[n,m]\} = \mathcal{A}(\zeta,\gamma) = \sum_{m=0}^{M-1}\sum_{m=0}^{N-1} a[n,m]\zeta^n\gamma^m, \tag{2.69}$$

where $\boldsymbol{Z}\{\cdot\}$ denotes the Z transform operation, (ζ,γ) are the cartesian coordinates in Z space and the function $\mathcal{A}(\zeta,\gamma)$ is known as the Z spectrum of the discrete function. Convergence of the summation in Eqn.(2.69) is guaranteed for all values of ζ and γ in the ranges $0 \leq |\zeta| < \infty$ and $0 \leq |\gamma| < \infty$ since the image is compact and this implies $|a[n,m]| < \infty$, $\forall\ n$ and m (Oppenheim and Schafer, 1975).

The Z transform is a unique invertible mapping (Siebert, 1986) and this mapping is denoted by the Z transform pair

$$a[n,m] \overset{ZT}{\Longleftrightarrow} \mathcal{A}(\zeta,\gamma).$$

It is possible to use various inversion techniques to obtain the discrete function from its Z transform or sampled Z transform. Techniques for inverting the Z transform are contained in many texts on digital signal processing (Oppenheim and Schafer, 1975; Siebert, 1986; Nagrath and Gopal, 1982), whilst techniques for reconstructing the discrete function from samples of

its Z transform are discussed in §4.5. It should be noted that Eqn.(2.69) is also in the form of a 2-D polynomial of degree $(N-1)(M-1)$ in the complex variables ζ and γ. This observation is particularly valuable as it implies that some of the properties of the Z transform can be inferred from the established properties of algebraic polynomials (Requicha, 1980).

The similarity between the Z transform and the Fourier transform can be further elucidated if the expressions for the Z transform and Fourier transform of a discrete image are compared in the manner outlined below. In order to make the comparison between the two expressions, it is first necessary to analytically continue the Fourier space variables, u and v, so that they can take on complex values, i.e,

$$\begin{aligned} u &= \Re(u) + i\Im(u), \\ v &= \Re(v) + i\Im(v), \end{aligned}$$

where $\Im(u)$ and $\Im(v)$ are not zero valued everywhere. Using these analytically continued Fourier space variables it is now possible to define the mappings

$$\begin{aligned} \zeta &= \exp^{i2\pi u\epsilon} \qquad (2.70) \\ &= \exp^{-2\pi\Im(u)}\exp^{i2\pi\Re(u)}, \\ \gamma &= \exp^{i2\pi v\epsilon} \qquad (2.71) \\ &= \exp^{-2\pi\Im(v)}\exp^{i2\pi\Re(v)}, \end{aligned}$$

which enable points in Z space to be mapped to corresponding points in Fourier space. On substituting the above mapping relationships into the expression for the Fourier transform of a discrete image (Eqn.(2.62)) it can be seen that the resulting expression and the expression for the 2-D Z transform (Eqn.(2.69)) are equivalent. Hence, this comparison shows that it is possible to map the Z transform and continuous Fourier transform of a discrete image directly between Z space and Fourier space.

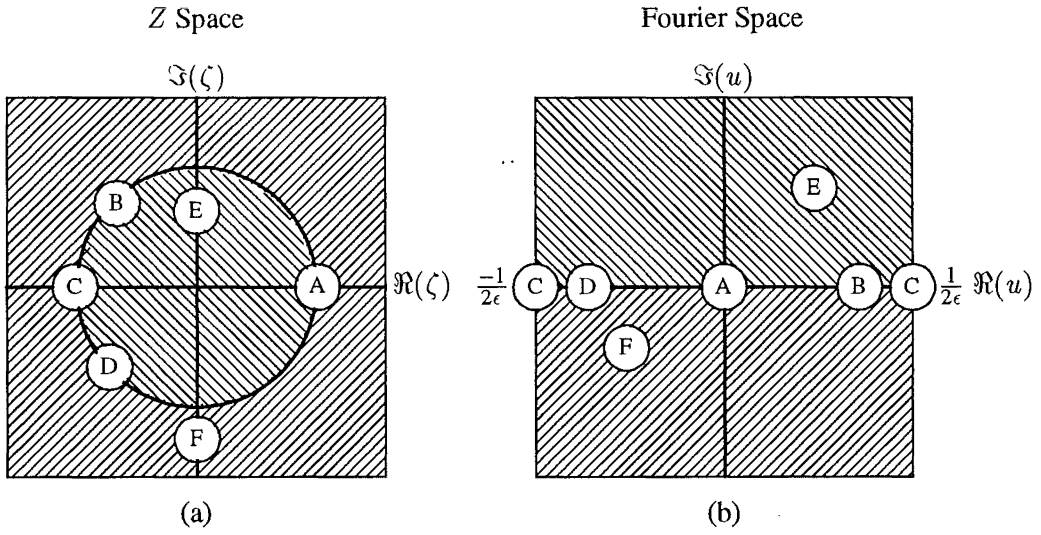

Figure 2.11 A demonstration of how the mappings $\zeta = \exp^{i2\pi u\epsilon}$ and $\gamma = \exp^{i2\pi v\epsilon}$ relate points in Fourier and Z space.

On evaluating the mapping relationships in Eqns. (2.70) and (2.71) for various values of u and v it is possible to derive a set of figures, such as those comprising Fig. 2.11, which show

how points in Z space map into Fourier space and vice versa. Fig. 2.11 shows that points on the ζ plane in Z space within the unit circle, on the unit circle, and outside the unit circle map to points in Fourier space on the upper half of the u plane, $\Re(u)$ axis, and lower half of the u plane respectively. It can also be seen that mapping the Z transform into Fourier space also results in a spectrum that is periodic in $\Re(u)$ with a period of $1/\epsilon$ which corresponds to one complete traversal of the unit circle in Z space. The periodic nature of the spectrum in Fourier space should be expected from the discussion in §2.5.3 where it is shown that a discrete (or equivalently sampled) image's spectrum must be periodic in Fourier space due to the convolution that occurs in image space when an image is sampled.

Z TRANSFORM PROPERTIES

Property	Spatial Domain		FrequencyDomain
Linearity	$k_1 r[\mathbf{n}] + k_2 q[\mathbf{n}]$	$\overset{ZT}{\Longleftrightarrow}$	$k_1 \mathcal{R}(\zeta) + k_2 Q(\zeta)$
Scaling	${k_1}^{-\mathbf{n}}\, r[\mathbf{n}]$	$\overset{ZT}{\Longleftrightarrow}$	$\mathcal{R}(\frac{\zeta}{k_1})$
Shifting	$r[\mathbf{x} - \mathbf{n}_r]$	$\overset{ZT}{\Longleftrightarrow}$	$\zeta^{-\mathbf{n}_r}\mathcal{R}(\zeta)$
Conjugate Functions	$r^*[\mathbf{n}]$	$\overset{ZT}{\Longleftrightarrow}$	$\mathcal{R}^*(\zeta^*)$
Reflection	$r[-\mathbf{n}]$	$\overset{ZT}{\Longleftrightarrow}$	$\mathcal{R}(\frac{1}{\zeta})\zeta^{N-1}$
Differentiation of $\mathcal{R}(\zeta)$	$\mathbf{n}r(\mathbf{n})$	$\overset{ZT}{\Longleftrightarrow}$	$\frac{d}{d\zeta}\mathcal{R}(\zeta)$
Convolution	$r[\mathbf{n}] \odot q[\mathbf{n}]$	$\overset{ZT}{\Longleftrightarrow}$	$\mathcal{R}(\zeta)\mathcal{Q}(\zeta)$
Cross-correlation	$r[\mathbf{n}] \star q[\mathbf{n}]$	$\overset{ZT}{\Longleftrightarrow}$	$\mathcal{R}^*(\frac{1}{\zeta^*})\mathcal{Q}(\zeta)\zeta^{N-1}$
Auto-correlation	$r[\mathbf{n}] \star r[\mathbf{n}]$	$\overset{ZT}{\Longleftrightarrow}$	$\mathcal{R}^*(\frac{1}{\zeta^*})\mathcal{R}(\zeta)\zeta^{N-1}$
Real in Spatial Domain	$\Im\{r[\mathbf{n}]\} = 0$	$\overset{ZT}{\Longleftrightarrow}$	$\mathcal{R}(\zeta) = \mathcal{R}^*(\zeta^*)$

Table 2.2

As with the Fourier transform, there are a number of properties that relate either the form of a spatial domain function to the form of its Z transform or that relate a spatial domain operation to its equivalent Z domain operation. The properties of the Z transform, which can be found in most digital signal processing texts (Oppenheim and Schafer, 1975; Nagrath and Gopal, 1982; Siebert, 1986) are listed in Table 2.2. Table 2.2 is an adapted version of the table that appears in Oppenheim and Schafer (1975). As with the Fourier transform the definition used here for the Z transform differs from most texts which tend to define the Z transform with a minus sign in the exponent. For this reason the Z transform properties in Table 2.2 differ from most texts with regards to the placement of minus signs in the shifting, scaling and differentiation properties. In this table the important properties to note are those of linearity, conjugation, reflection, convolution, and correlation due to their relevance to the deconvolution and phase retrieval problems to be introduced in §3.1. The cross-correlation and auto-correlation properties did not appear in the original table in Oppenheim and Schafer (1975), but the auto-correlation property can be derived from the convolution property if the auto-correlation is expressed, as in Eqn.(2.56), in the form of a convolution, i.e.

$$z_{rr}[n] = r^*[n] \odot r[-n].$$

On applying the conjugate and reflection properties of the Z transform to the above equation,

the Z transform of the auto-correlation is given by the expression

$$z_{rr}[n] \overset{ZT}{\Longleftrightarrow} \mathcal{R}^*(\zeta^*)\mathcal{R}(\frac{1}{\zeta})\zeta^{N-1}.$$

Replacing ζ by $1/\zeta$ gives

$$z_{rr}[n] \overset{ZT}{\Longleftrightarrow} \mathcal{R}^*(\frac{1}{\zeta^*})\mathcal{R}(\zeta)\zeta^{N-1},$$

which should be compared to the expression in Table 2.2 for the cross-correlation which was derived from first principles.

2.8 ENTIRE FUNCTIONS

Functions which can be categorised into the class of functions known as entire functions have many useful properties. The properties of entire functions have been used extensively in developing the deconvolution and phase retrieval algorithms contained within Chapters 5 and 6.

A function $C(w)$, which exists in the complex w plane, is said to be entire if $C(w)$ does not possess any singularities (i.e. is analytic) for all finite values of w (Requicha, 1980). In order for a complex function to be analytic at a point w_0 on the complex w plane it must satisfy the Cauchy-Riemann equations in the neighbourhood of w_0 (Requicha, 1980). Louiville's theorem states that if a general function $C'(w)$ is analytic and bounded within the finite w plane then $C'(w)$ is a constant (Kreyszig, 1983). In general, however, entire functions are not constant and therefore they must have singularities at infinity to satisfy Louiville's theorem. One particularly useful property of entire functions is that if one member of a Fourier transform pair is bounded (i.e. is of finite bandwidth or finite spatial extent) then the other member of the pair is an entire function (Requicha, 1980). For this reason in many practical situations the expressions used to model a large number of physical processes are themselves entire functions or can be transformed into entire functions.

A subclass of entire functions are the entire functions of exponential type (EFET). It is customary to classify the rate of growth of a function by comparing its rate of growth to a function of the form $\exp^{r^k}$ (Requicha, 1980). A function which grows no faster than $\exp^r$ (i.e. $k = 1$) is known as an entire function of exponential type. Examples of two types of function which can be classified as EFET are the algebraic and trigonometric polynomials. In the situation where a function is bounded and is of finite energy the other member of the Fourier transform pair comprising this function is an EFET (Requicha, 1980). It can therefore be seen that both the Z transforms and Fourier transforms of the compact images dealt with in this thesis are EFET. An important property of EFET is that the function can be represented, apart from an arbitrary constant complex scaling factor, by the set of zeros belonging to the function (Requicha, 1980). The set of zeros of the EFET $C(w)$ are defined to be the set of points on the w plane $\{w_0, w_1, w_2, \ldots\}$ where $C(w) = 0$. Thus, if a set of zeros of an EFET have been obtained by some means, it is possible to uniquely reconstruct the function characterised by these zeros to within the arbitrary complex constant. This property of EFET provides the foundations for the deconvolution and phase retrieval algorithms presented later in this thesis.

CHAPTER 3

DECONVOLUTION AND PHASE RETRIEVAL TECHNIQUES

This chapter introduces a number of different techniques which attempt to invert either the equation describing a blurred contaminated image (Eqn.(2.49)) or the equation describing an image's contaminated power spectrum (a contaminated version of Eqn.(2.58)) in an attempt to obtain an estimate of the original image. This review of deblurring (deconvolution) and phase retrieval algorithms aims to provide a framework within which both the new deconvolution and phase retrieval algorithms described in §5.1 and §6.1.2 fit. The review reveals that the existing 2-D deconvolution algorithms suffer from the limitation that the accuracy of their solutions is heavily dependent on the accuracy of the psf estimate used to obtain the solution. The new deconvolution algorithm described in §5.1 was proposed in the expectation that it would not suffer from this limitation. The deconvolution techniques described are limited in the main to those which are applied in the frequency domain without iteration and which use an estimate of the psf's spectrum to obtain an estimate of the image. The review concentrates on these types of deconvolution techniques because the new algorithm introduced in §5.1 is of this type. Likewise, the description of existing phase retrieval algorithms is primarily limited to iterative frequency domain techniques, since these are relevant to the new phase retrieval algorithm in §6.1.2.

In a situation requiring deconvolution it is likely that there is some prior information available which can be used to assist the deconvolution algorithm in obtaining a solution. As there are specific types of prior information available it is possible to categorise the assortment of deconvolution problems based on the available prior information. In the first section of this chapter the different types of deconvolution problem and the prior knowledge which characterises each of the problems are discussed. In §3.2 the inverse filter, Wiener filter, and 1-D zero-based consistent deconvolution algorithm, which are the existing algorithms most closely related to the new deconvolution algorithm in §5.1, are described. Also included within §3.2 is a brief review of some of the more common iterative and recursive deconvolution techniques which can also be used to solve the same deconvolution problem as the techniques above. A description of some methods which could be used to obtain a psf estimate is then given in §3.3. Of all the deconvolution techniques described in §3.2, the only technique which can provide what is known as a consistent solution to the uncontaminated deconvolution problem is the 1-D consistent deconvolution algorithm. The new algorithm in §5.1 also aims to provide a consistent solution to the deconvolution problem, consequently the principles of consistent deconvolution are explained and an explanation of why a consistent solution is so desirable is given in §3.4. To complete this chapter, the iterative algorithms which are commonly used to solve the phase retrieval problem and conditions necessary to guarantee a unique solution are described in §3.5.

3.1 CATEGORIES OF DECONVOLUTION PROBLEM

The deconvolution problem, particularly in the situation where the convolution has been corrupted by any form of contamination, is a prime example of a ill-conditioned inverse problem (Andrews and Hunt, 1977; Katsaggelos, 1989; Biemond *et al.*, 1990; Katsaggelos, 1991). To obtain a useful solution to any inverse problem which is ill-conditioned it is necessary to have some prior information about the form of the solution or process of interest. This prior information enables the solution to be constrained in some manner and the solution process to be stabilised. In all deconvolution problems some form of prior knowledge on the form of the blurring, original image, or type of contamination is available, otherwise there would be no basis for stating that the recorded image was degraded. The exact nature of this prior information dictates to the largest degree which type of deconvolution algorithm must be used to obtain an estimate of the original image. It is, therefore, possible to categorise all deconvolution problems based on the available prior information. The different types of deconvolution problem, and the form of the prior knowledge that characterises each problem, are now described.

The first deconvolution problem to be described, and that which the new algorithm in this thesis aims to solve, is known as the standard deconvolution problem. The feature that distinguishes the standard deconvolution problem from the other problems to be described is that an estimate of the psf is available *a priori*. The standard deconvolution problem can therefore be posed: given $g(\mathbf{x})$ and an estimate of $h(\mathbf{x})$, recover an estimate of $f(\mathbf{x})$ (Bates and McDonnell, 1986; Bates *et al.*, 1990). This form of deconvolution problem is quite common in practice, since it is often possible to estimate $h(\mathbf{x})$ from the physics of the imaging system or from $g(\mathbf{x})$ using one of the psf estimation techniques described in §3.3. Well known techniques that can be used to solve the standard deconvolution problem, including inverse, Wiener, geometrical mean and constrained least squares filtering, are described in §3.2. Some of the techniques described, such as constrained least squares filtering, can also use prior information in addition to the psf estimate to further constrain the form of the solution.

In many scientific and technological imaging situations an image of an object is formed by the passage of electromagnetic radiation through a turbulent medium, e.g. in the case of astronomical imaging the turbulent medium is earth's atmosphere. In these imaging situations it is often possible to obtain a collection of images of the same object and, if the time between obtaining each image is long enough and the exposures are short enough, the contamination and blurring on each image should be statistically independent (Satherley and Bones, 1993). If this is the case, algorithms have been developed which can reconstruct the common feature in each of the images (i.e. the image of the object) by using an averaging process which removes the degradation by destructively superimposing the blurring and contamination in the images. This type of deconvolution problem is known as the ensemble blind deconvolution problem and can be stated as: given M different versions of $g(\mathbf{x})$, i.e.

$$g_m(\mathbf{x}) = f(\mathbf{x}) \odot h_m(\mathbf{x}) + c_m(\mathbf{x}), m = 1, 2, \ldots, M,$$

where the $h_m(\mathbf{x})$ and $c_m(\mathbf{x})$ are statistically independent, obtain an estimate of $f(\mathbf{x})$ (Bates *et al.*, 1990; Satherley and Bones, 1993). It should be noted that with the ensemble blind deconvolution problem no prior information other than the statistically independent nature of the $h_m(\mathbf{x})$ and $c_m(\mathbf{x})$ is available, i.e. the $\hat{h}_m(\mathbf{x})$ are not known. There are only a limited number of techniques available which can be used to solve this particular deconvolution problem since most of the techniques developed to solve the ensemble deconvolution problem in the astronomical arena, such as the algorithms of Labeyrie (1970), Knox and Thompson (1974) and Lohmann

et al. (1983), all require psf estimates. Techniques that can solve the ensemble blind deconvolution problem directly include the straightforward shift-and-add algorithm (Bates, 1976), zero-and-add algorithm (Davey *et al.*, 1986) and zero track (see §4.2.1) recognition algorithm (Satherley and Bones, 1993).

The final type of deconvolution problem to be considered here is the pure blind deconvolution problem. This problem can be stated: given a single blurred and contaminated image $g(\mathbf{x})$, in the situation where no information on $h(\mathbf{x})$ is known, obtain an estimate of $f(\mathbf{x})$ (Bates *et al.*, 1990). Another problem which is also in effect a pure blind deconvolution problem is the phase retrieval problem since this problem can be posed as given $|F(\mathbf{u})|^2$ (see Eqn.(2.58)) obtain an estimate of $f(\mathbf{x})$. The pure blind deconvolution problem can only be accurately solved in the case where $g(\mathbf{x})$ or $|F(\mathbf{u})|^2$ are uncontaminated if no prior information is available on the form of the solution. Thus, the phase retrieval algorithm of Lane *et al.* (1987), described in §4.3.2, cannot obtain a useful estimate of the image when $|F(\mathbf{u})|^2$ is contaminated and the recently reported blind deconvolution algorithm of Ghiglia *et al.* (1993), in its present form, would also be unable to provide a solution in the presence of contamination. In most practical blind deconvolution situations it is likely that some from of prior information on the form of the solution or contamination would be available to assist in the procurement of a solution. The prior information which could be available to assist an algorithm in obtaining a solution are estimates of the image's support and energy, whether the image is positive or complex, and the contamination's energy. There are a number of algorithms that use the above types of information to obtain a solution to the blind deconvolution or phase retrieval problems. These include: the method of projection onto convex sets (Sezan and Stark, 1982; Youla and Webb, 1982; Youla, 1987), the iterative phase retrieval algorithms (see §3.5) (Fienup, 1982; Bates and McDonnell, 1986; Bates and Mnyama, 1986), maximum entropy (Gull and Skilling, 1984; Newman and Hildebrandt, 1987), and homomorphic filtering (Newman and Hildebrandt, 1987).

3.2 SOLUTIONS TO THE STANDARD DECONVOLUTION PROBLEM

In this section the various methods which can be used to solve the standard deconvolution problem, as defined in §3.1, are described. This description comprises a reasonably detailed overview of the direct techniques which obtain solutions to the standard deconvolution problem by transforming the problem into Fourier space. To complete the review a brief overview of the most common iterative and recursive standard deconvolution techniques is also given. A summary of the deconvolution techniques discussed along with their advantages and disadvantages is given in Table 3.1

3.2.1 Inverse Filtering

The form of the spectrum of an uncontaminated blurred image $b(\mathbf{x})$, i.e. the product of the spectra of the original image and psf (see Eqn.(2.52)), suggests that a blurred image could be straightforwardly deblurred by dividing $B(\mathbf{u})$ by an estimate of the psf's spectrum and transforming back to image space (Andrews and Hunt, 1977; Bates and McDonnell, 1986; Biemond *et al.*, 1990). This deconvolution technique is known as inverse filtering because the filter, $\Xi_i(\mathbf{u})$, used to perform the deconvolution is defined by the expression

$$\Xi_i(\mathbf{u}) = \frac{1}{\hat{H}(\mathbf{u})}, \tag{3.1}$$

STANDARD DECONVOLUTION TECHNIQUES

Technique	Type Tech.	PSF Sens.	Strengths	Limitations
Inverse Filter.	D.	Y.	Straightforward. Reconstructs high frequencies.	Sensitive to contamination. Divide by zero problems. High frequency artifacts.
Wiener Filter.	D.	Y.	Robust w.r.t. high frequency contamination.	Robustness obtained at expense of resolution. Usually applied iteratively.
Constrained Least Squares.	D.	Y.	Incorporates prior information on contamination level and removes high frequency artifacts.	Removes high frequency artifacts at the expense of deblurring effectiveness.
Geometrical Mean Filter.	D.	Y.	Eliminates undesirable properties of inverse and Wiener filter.	Solution quality dependent on parameter η.
POCs.	I.	Y.	Constrains solution using additional prior information.	Not all prior information forms a convex set.
Van Cittert Iteration.	I.	Y.	See inverse filter. Iteration can be stopped before inverse filter problems arise. Can incorporate other prior information.	See inverse filter.
Kalman Filtering.	R.	Y.		Image and psf models must be determined.
Maximum Entropy.	I.	Y.	Can incorporate additional prior information. Solution is maximally noncommittal about unknown characteristics, while satisfying prior information.	Generally computationally expensive.
1-D consistent.	D.	N.	Insensitive to psf inaccuracy.	Only 1-D images can be deconvolved.

Table 3.1 Key: PSF Sens. = Is solution sensitive to psf estimate accuracy?, Type Tech. = Type of Technique, Y. = Yes, N. = No, D. = Direct., I. = Iterative, and R. = Recursive.

where $\hat{H}(\mathbf{u})$ is an estimate of the psf's spectrum. In the uncontaminated situation the inverse filter provides an optimum solution to the deconvolution problem, in the sense that it minimises the squared error between the image estimate and the original image (Andrews and Hunt, 1977), i.e. the inverse filter minimises

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left|B(\mathbf{u})-[\Xi_i(\mathbf{u})]^{-1}\hat{F}(\mathbf{u})\right|^2 d\mathbf{u}. \tag{3.2}$$

On expressing the spectrum of the image estimate, provided by inverse filtering, in the form

$$\hat{F}(\mathbf{u}) = B(\mathbf{u})\Xi_i(\mathbf{u}),$$

one of the problems with the inverse filter immediately becomes apparent. This problem is that the quality of the image estimate is directly related to the accuracy of the psf estimate and as a consequence $\hat{f}(\mathbf{x})$, while it is a least squares solution, is not a consistent solution to the deconvolution problem (see §3.4). Another problem with the inverse filtering technique is that an inverse filter cannot be defined for all $\hat{H}(\mathbf{u})$. This is particularly true for those $\hat{H}(\mathbf{u})$ that have zeros within the regions of Fourier space where $B(\mathbf{u})$ is non-zero, since the filter is undefined at those values of $\mathbf{u}$ where $\hat{H}(\mathbf{u})$ is zero. The spectra of many common types of psf, such as out-of-focus and linear motion psfs, are likely to have many zeros within the region of Fourier space where $B(\mathbf{u})$ is non-zero and consequently images blurred by these types of psf cannot be successfully deblurred by inverse filtering techniques.

In the situation where the blurred image has also been contaminated, the inverse filter produces an image estimate whose spectrum is defined by the expression

$$\begin{aligned}\hat{F}(\mathbf{u}) &= \Xi_i(\mathbf{u})[F(\mathbf{u})H(\mathbf{u})+C(\mathbf{u})]\\ &\approx F(\mathbf{u})+\Xi_i C(\mathbf{u})\end{aligned} \tag{3.3}$$

Most practical images have a spectrum where the majority of the energy is confined to the lower frequency components and the energy steadily decreases with increasing frequency. In contrast, the spectra of typical contamination functions usually have quite significant frequency components at high frequencies (Biemond *et al.*, 1990). This is also often true of $\Xi_i(\mathbf{u})$ since many common psfs (such as out-of-focus, linear motion, and Gaussian blurring) have a spectral magnitude that decreases rapidly with increasing frequency. Thus, Eqn.(3.3) demonstrates that for a typical blurred contaminated image the image estimate provided by the inverse filter is likely to have significant spectral components at high spatial frequencies that consist almost entirely of inverse filtered contamination. In the regions of the spectrum where $\hat{H}(\mathbf{u})$ tends to zero it is also very likely that the term $\Xi_i C(\mathbf{u})$ will be much larger than $F(\mathbf{u})$ and the most significant contribution to $\hat{F}(\mathbf{u})$ is again the inverse filtered contamination. The inverse filter is therefore often unable to provide a useful solution to the contaminated deconvolution problem in practice.

A demonstration of the ability of the inverse filter to provide a solution to the uncontaminated and contaminated deconvolution problems is shown in Fig. 3.1. The reconstructions obtained on inverse filtering a contaminated and uncontaminated version of the blurred image in Fig. 3.1 (b) are shown in Fig. 3.1 (c) and Fig. 3.1 (d) respectively. The uncontaminated blurred image was formed by convolving the original 64×64 image in Fig. 3.1 (a) with a bilevel psf of extent 5 pixels. The psf consisted of a 4×4 pixel square centrally situated on a lower intensity 5×5 pixel background. The contaminated image was produced by adding Gaussian distributed noise to the uncontaminated blurred image to give the resulting image a 60 dB SNR. On comparing the reconstructions in Fig. 3.1 (c) and Fig. 3.1 (d) to the original image

in Fig. 3.1 (a) it can be seen that the reconstruction from the contaminated convolution does indeed contain significant high frequency artifacts. The quality factors for the reconstructions are Q=86 dB and Q=8 dB for the uncontaminated and contaminated convolutions respectively. The reconstruction in Fig. 3.1 (d) demonstrates that even with a high convolution SNR the inverse filter is unable to obtain a particularly useful reconstruction.

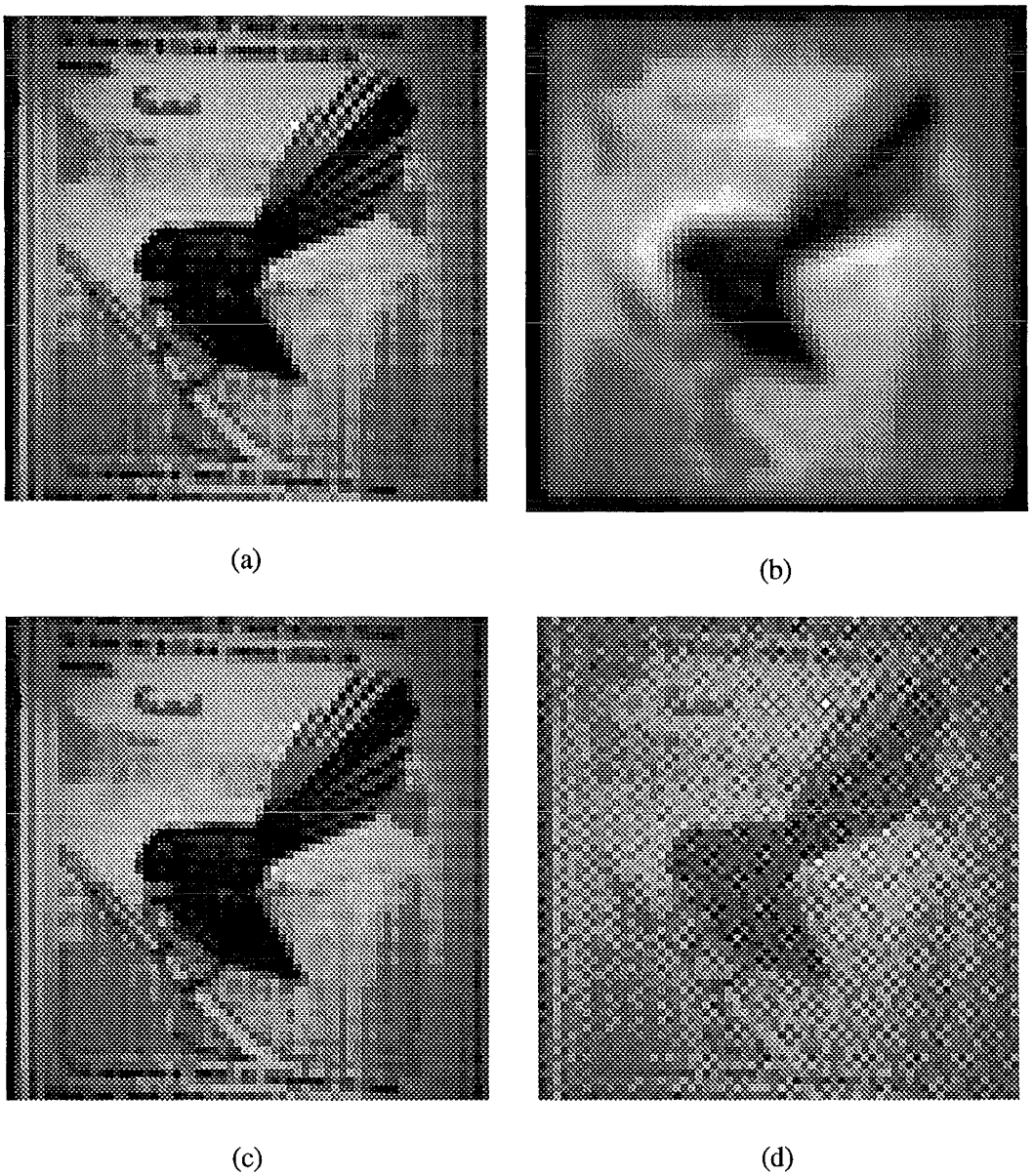

Figure 3.1 The original 64×64 pixel image in (a) has been blurred using a bilevel 5×5 psf approximating a disk to give the blurred image in (b). The reconstructed images in (c) and (d) were reconstructed using an inverse filter from a uncontaminated and 60 dB SNR contaminated version of (b) respectively.

On performing a more in-depth evaluation of the contaminated deconvolution problem and the limitations of inverse filtering it is possible to derive two more sophisticated filters, based on the inverse filter, that can provide more accurate estimates of the image in the contaminated situation. The principles surrounding the development of the first of these filters are based on the recognition that the inverse filter attempts to provide a $\hat{f}(\mathbf{x})$, that if it were to be

reblurred, would provide an exact replica of the observed blurred and contaminated image $g(\mathbf{x})$. Consequently, the inverse filter unreasonably attempts to find an $\hat{f}(\mathbf{x})$ that on reblurring provides a more accurate representation of the observed image than that provided by $f(\mathbf{x})$ (Biemond *et al.*, 1990). Instead of attempting to find an exact match, as the inverse filter attempts, it is far more reasonable to constrain the form of the final solution by using an inverse type of filter that also makes use of prior knowledge regarding the magnitude of the contamination and limitations of the inverse filter.

The constrained least squares filter is a filter that provides more accurate reconstructions by imposing two constraints on the form of the solution. The first of these constraints is the more reasonable expectation that $\hat{F}(\mathbf{u})$ should satisfy the relationship

$$\|G(\mathbf{u}) - \hat{F}(\mathbf{u})H(\mathbf{u})\| = \|C(\mathbf{u})\|, \tag{3.4}$$

where $\|C(\mathbf{u})\|$ denotes the Euclidean norm $\sqrt{(K)\int |C(\mathbf{u})|^2 d\mathbf{u}}$ and similarly for other functions, and the second criterion is that $\hat{F}(\mathbf{u})$ is a quantity that minimises the expression

$$\|Q(\mathbf{u})\hat{F}(\mathbf{u})\|,$$

where $Q(\mathbf{u})$ is the transfer function of an operator that measures the smoothness of the reconstruction (Andrews and Hunt, 1977; Biemond *et al.*, 1990). An estimate of $\|C(\mathbf{u})\|$ can usually be obtained from a smooth region in the image (Biemond *et al.*, 1990). The second constraint aims to minimise the high frequency artifacts known to occur with inverse filtering and the function $Q(\mathbf{u})$, used to measure the magnitude of these artifacts, is generally a second derivative operator such as a 2-D Laplace Filter (Biemond *et al.*, 1990). On solving the above minimisation problem the constrained least squares filter, defined by the transfer function

$$\Xi_{cl}(\mathbf{u}) = \frac{\hat{H}^*(\mathbf{u})}{|\hat{H}(\mathbf{u})|^2 + \alpha|Q(\mathbf{u})|^2}, \tag{3.5}$$

is obtained (Andrews and Hunt, 1977; Biemond *et al.*, 1990). In Eqn.(3.5) $\frac{1}{\alpha}$ is a Lagrange multiplier chosen so that Eqn.(3.4) is satisfied. Eqn.(3.5) demonstrates that the quality of the reconstructions are directly related to the quality of the psf estimate. It is worth noting that in the uncontaminated situation α must be zero, reducing the constrained least squares filter to the inverse filter, otherwise Eqn.(3.4) would not be satisfied.

The form of the second type of filter that can provide improved reconstructions is found by minimising a squared error expression, similar to Eqn.(3.2), in the contaminated situation. The filter provided by performing this minimisation is the least squares or Wiener filter which is discussed in the following section.

3.2.2 Wiener Filter

It was shown in §3.2.1 that the most straightforward deconvolution technique, the inverse filter, is unlikely to be able to provide a useful solution to the contaminated deconvolution problem. A filter which can provide a useful reconstruction in the contaminated situation is the Wiener filter, $\Xi_w(\mathbf{u})$, which aims to minimise the square error between an original image and the image estimate obtained from a contaminated convolution of the image. In this section the form of the Wiener filter is derived from this minimum square error criterion. The primary motivation for presenting the complete derivation of the Wiener filter here is that the Wiener filter is the deconvolution technique with which the performance of the new algorithm in §5.1

is compared and, consequently, a complete understanding of the form of this filter and its limitations is desirable. The derivation of $\Xi_w(\mathbf{u})$ presented here follows the derivation of Rosenfeld and Kak (1982), as this derivation fits most comfortably within this thesis and is considered, by the author, to be the most straightforward derivation. The interested reader can find other derivations of $\Xi_w(\mathbf{u})$ from the minimum squared error criterion in Andrews and Hunt (1977) and Castleman (1979). Also included within this section is a brief description of the geometrical mean filter which combines both the Wiener and inverse filter in an attempt to overcome the limitations of both filters.

The aim of the derivation that follows is to deduce the form of the filter function that can produce an image estimate from $g(\mathbf{x})$ that minimises the mean error expression

$$e^2 = E\{[f(\mathbf{x}) - \hat{f}(\mathbf{x})]^2\}, \tag{3.6}$$

In Eqn.(3.6) e^2 denotes the mean square error and E is the expected value operator. If it is now assumed that the filter which is to perform the deconvolution is a multiplicative filter, then the image estimate can be expressed in the form of the convolution

$$\hat{f}(\mathbf{x}) = \int \xi_w(\mathbf{x} - \mathbf{x}')g(\mathbf{x}')d\mathbf{x}',$$

and Eqn.(3.6) can in turn be expressed in the form

$$e^2 = E\left\{\left[f(\mathbf{x}) - \int \xi_w(\mathbf{x} - \mathbf{x}')g(\mathbf{x}')d\mathbf{x}'\right]^2\right\}, \tag{3.7}$$

where $\xi_w(\mathbf{x})$ is the filter function to be determined. It is now shown, using the method of Rosenfeld and Kak (1982), that the filter function that minimises Eqn.(3.7) also satisfies the expression

$$E\left\{\left[f(\mathbf{x}) - \int \xi_w(\mathbf{x} - \mathbf{x}')g(\mathbf{x}')d\mathbf{x}'\right] g(\mathbf{s})\right\} = 0, \tag{3.8}$$

where $\mathbf{s}$ is a position vector in image space.

To demonstrate that the $\xi_w(\mathbf{x})$ that satisfies Eqn.(3.8) also minimises Eqn.(3.7), consider the two filter functions $\xi_w(\mathbf{x})$ and $\hat{\xi}_w(\mathbf{x})$ which satisfy and do not satisfy Eqn.(3.8) respectively. If the filter $\hat{\xi}_w(\mathbf{x})$ is used to perform the deconvolution the mean square error in the solution is given by

$$\begin{aligned}
\hat{e}^2 &= E\left\{\left[f(\mathbf{x}) - \int \hat{\xi}_w(\mathbf{x} - \mathbf{x}')g(\mathbf{x}')d\mathbf{x}' - \int \xi_w(\mathbf{x} - \mathbf{x}')g(\mathbf{x}')d\mathbf{x}' + \int \xi_w(\mathbf{x} - \mathbf{x}')g(\mathbf{x}')d\mathbf{x}'\right]^2\right\}, \\
&= E\left\{\left[f(\mathbf{x}) - \int \xi_w(\mathbf{x} - \mathbf{x}')g(\mathbf{x}')d\mathbf{x}' + \int [\xi_w(\mathbf{x} - \mathbf{x}') - \hat{\xi}_w(\mathbf{x} - \mathbf{x}')]g(\mathbf{x}')d\mathbf{x}'\right]^2\right\}, \\
&= e^2 + E\left\{\left[\int [\xi_w(\mathbf{x} - \mathbf{x}') - \hat{\xi}_w(\mathbf{x} - \mathbf{x}')]g(\mathbf{x}')d\mathbf{x}'\right]^2\right\} \\
&\quad +2E\left\{\left[f(\mathbf{x}) - \int \xi_w(\mathbf{x} - \mathbf{x}')g(\mathbf{x}')d\mathbf{x}'\right]\left[\int [\xi_w(\mathbf{x} - \mathbf{x}') - \hat{\xi}_w(\mathbf{x} - \mathbf{x}')]g(\mathbf{x}')d\mathbf{x}'\right]\right\}
\end{aligned} \tag{3.9}$$

On changing the variable of integration in the second factor comprising the last term of Eqn.(3.9) from $\mathbf{x}'$ to $\mathbf{s}$ and interchanging the order of expectation and integration, $\hat{e}^2$ can be expressed in

the form

$$\begin{aligned}\hat{e}^2 &= e^2 + E\left\{\left[\int[\xi_w(\mathbf{x}-\mathbf{x}') - \hat{\xi}_w(\mathbf{x}-\mathbf{x}')]g(\mathbf{x}')d\mathbf{x}'\right]^2\right\} \\ &\quad +2\int E\left\{\left[f(\mathbf{x}) - \int \xi_w(\mathbf{x}-\mathbf{x}')g(\mathbf{x}')d\mathbf{x}'\right] g(\mathbf{s})\right\}\left[[\xi_w(\mathbf{x}-\mathbf{s}) - \hat{\xi}_w(\mathbf{x}-\mathbf{s})]\right] d\mathbf{s},\end{aligned}$$

which can be reduced to

$$\hat{e}^2 = e^2 + E\left\{\left[\int[\xi_w(\mathbf{x}) - \mathbf{x}') - \hat{\xi}_w(\mathbf{x}-\mathbf{x}')]g(\mathbf{x}')d\mathbf{x}'\right]^2\right\}, \tag{3.10}$$

since $\xi_w(\mathbf{x})$ satisfies Eqn.(3.8). The second term in Eqn.(3.10) must always be positive, which implies that $\hat{e}^2 \geq e^2$, and therefore the $\xi_w(\mathbf{x})$ satisfying Eqn.(3.8) does indeed provide the solution with the minimum e^2.

As it is now known that the $\xi_w(\mathbf{x})$ satisfying Eqn.(3.8) provides the solution with the least squared error the exact form of $\xi_w(\mathbf{x})$ can be determined from Eqn.(3.8). On rearranging Eqn.(3.8) and interchanging the order of expectation and integration the expression

$$\int\int \xi_w(\mathbf{x}-\mathbf{x}')E\{g(\mathbf{x}')g(\mathbf{s})\}d\mathbf{x}' = E\{f(\mathbf{x})g(\mathbf{s})\}, \tag{3.11}$$

is obtained. For a LTI system $E\{g(\mathbf{x}')g(\mathbf{s})\}$ and $E\{f(\mathbf{x})g(\mathbf{s})\}$ are equivalent to the correlation functions $z_{gg}(\mathbf{x}'-\mathbf{s})$ and $z_{fg}(\mathbf{x}-\mathbf{s})$ respectively (Bendat and Pierslo, 1980; Rosenfeld and Kak, 1982). If the substitutions $\mathbf{x}'-\mathbf{s} = \mathbf{t}$ and $\mathbf{x}-\mathbf{s} = \tau$ are used in Eqn.(3.11), the equation becomes

$$z_{fg}(\tau) = \int\int \xi_w(\mathbf{t}-\tau)z_{gg}(\mathbf{t})d\mathbf{t},$$

which can then be Fourier transformed to give the expression

$$\Xi_w(\mathbf{u}) = \frac{Z_{fg}(\mathbf{u})}{Z_{gg}(\mathbf{u})},$$

describing the form of the filter in terms of the spectral density of the convolution, $Z_{gg}(\mathbf{u})$, and the cross-spectral density of the convolution and image, $Z_{fg}(\mathbf{u})$. On assuming that $c(\mathbf{x})$ can be adequately described by a zero-mean Gaussian-distributed noise source which is uncorrelated with $f(\mathbf{x})$, it is possible to express the above spectral densities in the form

$$Z_{fg}(\mathbf{u}) = H^*(\mathbf{u})Z_{ff}(\mathbf{u}) \tag{3.12}$$

and

$$Z_{gg}(\mathbf{u}) = Z_{ff}(\mathbf{u})|H(\mathbf{u})|^2 + Z_{cc}(\mathbf{u}). \tag{3.13}$$

As a consequence the filter which can deconvolve a contaminated blurred image in the least mean squared error sense is given by

$$\Xi_w(\mathbf{u}) = \frac{H^*(\mathbf{u})}{|H(\mathbf{u})|^2 + \Phi(\mathbf{u})}, \tag{3.14}$$

where $\Phi(\mathbf{u}) = \frac{Z_{cc}(\mathbf{u})}{Z_{ff}(\mathbf{u})}$. This then is the Wiener filter and it can be seen that in the uncontaminated situation, where $Z_{cc}(\mathbf{u}) = 0$, the expression for the Wiener filter reduces to that describing the inverse filter (Eqn.(3.1)).

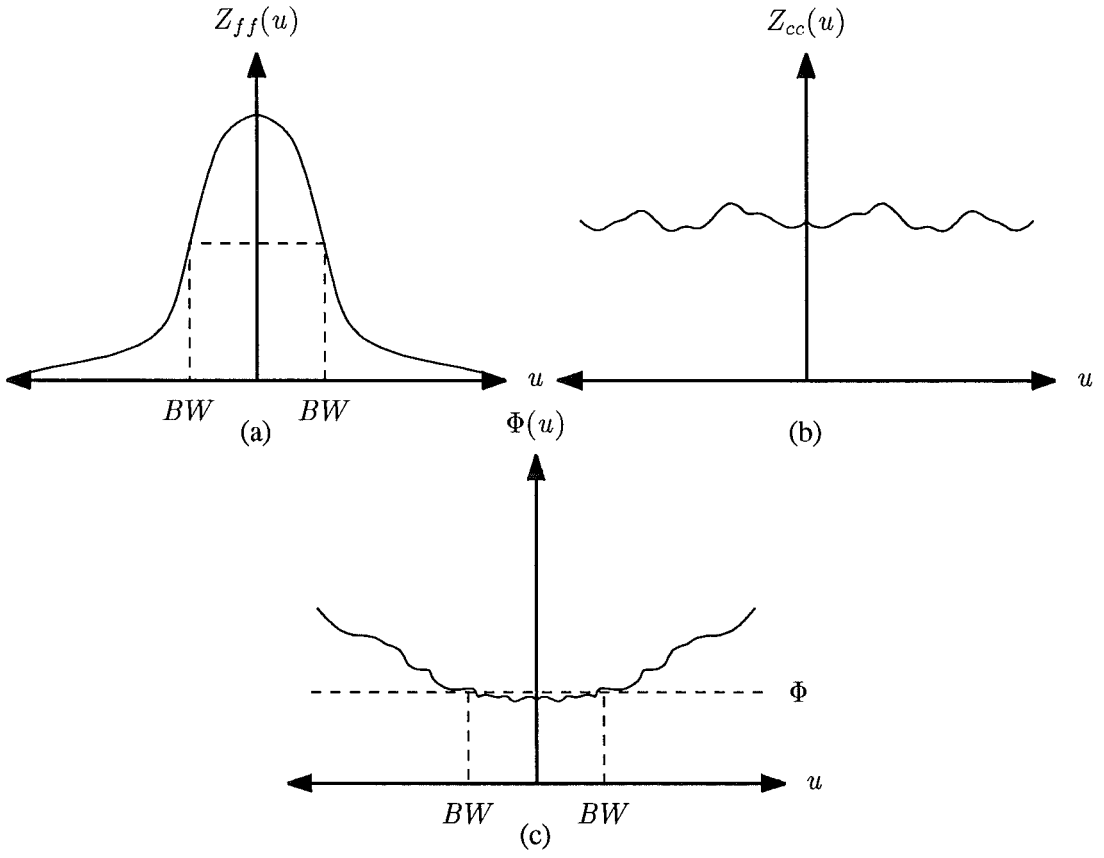

Figure 3.2 In (a) and (b) typical image and contamination power spectra are shown respectively. A typical $\Phi(\mathbf{u})$ derived from power spectra of the form of (a) and (b) is shown in (c).

In practice it is not usually possible to accurately estimate the spectral density functions of either the image or the contamination, consequently an approximation to $\Phi(\mathbf{u})$, which provides useful image estimates, is usually used. One of the most common approximation used for $\Phi(\mathbf{u})$ is to assume that it is equal to a constant Φ, known hereafter as the Wiener filter constant, which has a value that approximates the noise-to-signal ration of $g(\mathbf{x})$ (Rosenfeld and Kak, 1982; Bates and McDonnell, 1986). In order to obtain the best image estimate from a Wiener filter of this type, it is usual to vary Φ over a range of values and then subjectively judge which of the resulting $\hat{f}(\mathbf{x})$ is the best image estimate. The justification for using the approximation that $\Phi(\mathbf{u}) = \Phi$ is that the magnitude of the power spectrum of a typical image decreases rather rapidly with increasing frequency, whilst the power spectra of usual contamination functions tend to have a reasonably constant magnitude (i.e. white noise) or a magnitude that decreases at such a slow rate with frequency that the spectral density can be considered to be constant, i.e. the change in the magnitude of the spectral density function of the contamination occurs at a rate much slower than that of the image (Castleman, 1979). As a result $\Phi(\mathbf{u})$ calculated from these typical spectral density functions increases rather slowly with increasing frequency, for the values of $\mathbf{u}$ within the image's bandwidth, and $\Phi(\mathbf{u})$ in this region can be approximated by the constant Φ (see Fig. 3.2).

As an example of the ability of the Wiener filter to provide useful reconstructions, in the situation where $\Phi(\mathbf{u})$ has been approximated by Φ, the image estimate obtained on Wiener filtering the blurred and contaminated image in Fig. 3.1 (b) is shown in Fig. 3.3. The Wiener

filter used in this situation was constructed from the actual psf and the Wiener filter constant was set to 1.0×10^{-6}, which is the noise-to-signal ratio of the blurred contaminated image used in Fig. 3.1. On comparing the image estimate in Fig. 3.3 to that in Fig. 3.1 (d), obtained by inverse filter, it can be seen that the Wiener filter provides a reconstruction which is far superior to that obtained by inverse filtering (Q = 21 dB). The main reason for this superiority is that the Wiener filter constant has a stabilising effect on the filtering operation and reduces the magnitude of the filtered contamination component in $\hat{F}(\mathbf{u})$ at both high frequencies and as $H(\mathbf{u})$ tends to zero.

Figure 3.3 A reconstruction (Q = 21 dB) of the 60 dB SNR blurred image in Fig. 3.3 obtained by using a Wiener filter with $\Phi = 1.0 \times 10^{-6}$.

While the Wiener filter does not produce the filtered contamination artifacts that plague the reconstructions from the inverse filter, there remains the limitation that the quality of $\hat{f}(\mathbf{x})$ is highly dependent on the accuracy of $\hat{h}(\mathbf{x})$. The reader's attention is drawn to the above limitation of the Wiener filter as the new deconvolution technique, described in §5.1, is significantly less sensitive to the quality of the psf estimate than the Wiener filter. In addition to the above problem the Wiener filter also has a number of its own inherent limitations. The first of these is that the accuracy of the reconstructions is also highly dependent on the value chosen for the Wiener filter constant or the form of the function $\Phi(\mathbf{u})$. Furthermore, because it is usual to change the value of the Wiener filter constant until the best (subjectively judged) $\hat{f}(\mathbf{x})$ is obtained, the Wiener filter is in effect an iterative technique. Another major limitation of the Wiener filter is that a number of practical contamination sources, such as film grain noise and photon statistical noise on a CCD (see §2.3.2), are highly correlated with the original image and as a consequence the assumptions made in deriving Eqns. (3.12) and (3.13), required to determine the form of the Wiener filter, are not valid (Castleman, 1979). Thus, in the situation where a blurred image has been corrupted by the above forms of contamination it is very likely that the Wiener filter would be unable to provide an optimal estimate of the original image.

The final limitation of the Wiener filter, and a common criticism of its performance, is that the filter is only able to suppress the effects of the contamination at the expense of its ability to perform the deblurring operation (i.e. the Wiener filter is considered to perform only minor deblurring)(Biemond *et al.*, 1990). The reason for this behaviour is that $\Phi(\mathbf{u})$ attempts to suppress the high frequency contamination giving effectively a low pass filtered estimate of

$f(\mathbf{x})$ that is likely to have been filtered to a larger degree than the eye can tolerate (Andrews and Hunt, 1977). The geometrical mean filter, $\Xi_{gm}(\mathbf{u})$, overcomes this limitation of the Wiener filter by combining the desirable properties of both the Wiener and inverse filter in a filter defined by the transfer function

$$\Xi_{gm}(\mathbf{u}) = [\Xi_{inv}(\mathbf{u})]^{\eta}[\Xi_{w}(\mathbf{u})]^{1-\eta},$$

where $0 < \eta < 1$ (Andrews and Hunt, 1977). The inverse filter component in $\Xi_{gm}(\mathbf{u})$ allows $\hat{F}(\mathbf{u})$ to possess high frequencies, which the Wiener filter would otherwise suppress, and the Wiener filter component avoids the singularities associated with the zeros of $\hat{H}(\mathbf{u})$ which are an inherent problem with the inverse filter. Clearly this filter is also sensitive to inaccuracies in $\hat{H}(\mathbf{u})$ like its component filters. It is worth noting that Andrews and Hunt (1977) state that whilst the performance of the Wiener filter is a marked improvement on that of the inverse filter, its performance is poor compared to other filters such as the constrained least squares filter (see §3.2.1) and the geometric mean filter.

3.2.3 Other Standard Deconvolution Techniques

In this section a brief review of indirect techniques which can be used to solve the standard deconvolution problem are given. This review is by no means complete, as a thorough review is beyond the scope of this thesis and there are many variants of the techniques described here. This review and the previous section should, however, give the reader a reasonable overview of the more common standard deconvolution techniques and their limitations. For more thorough reviews of deconvolution techniques the interested reader is referred to the literature of Andrews and Hunt (1977), Biemond *et al.* (1990) and Katsaggelos (1991).

3.2.3.1 Van Cittert Iteration

The first method to be outlined here is an iterative spatial domain technique known as either the Van Cittert, Bially, or Landweber iteration method. Van Cittert's technique can be derived by expressing the original image in the form of the identity

$$f(\mathbf{x}) = f(\mathbf{x}) + \beta\left[b(\mathbf{x}) - h(\mathbf{x}) \odot f(\mathbf{x})\right], \tag{3.15}$$

which must hold for all values of the parameter β. On applying the method of successive substitutions to Eqn.(3.15) the iterative deconvolution scheme (Katsaggelos, 1989; Biemond *et al.*, 1990),

$$\begin{aligned}
\hat{f}_0(\mathbf{x}) &= \beta g(\mathbf{x}) \\
\hat{f}_{k+1}(\mathbf{x}) &= \hat{f}_k(\mathbf{x}) + \beta\left[g(\mathbf{x}) - h(\mathbf{x}) \odot \hat{f}_k(\mathbf{x})\right] \\
&= \beta g(\mathbf{x}) + [\delta(\mathbf{x}) - \beta h(\mathbf{x})] \odot \hat{f}_k(\mathbf{x}) \quad k = 1,2,\ldots,
\end{aligned}$$

is obtained. For the above iterative scheme to converge to a solution the norm $\|\delta(\mathbf{x}) - \beta h(\mathbf{x})\|$ must be less than unity (Katsaggelos, 1989; Biemond *et al.*, 1990). Biemond *et al.* (1990) show that in the limit as $k \to \infty$ the solution provided by Van Cittert iteration converges to the same solution which would be provided by the inverse filter. As a consequence, this iterative deconvolution scheme suffers from many of the limitations of the inverse filter. The advantages of this technique, however, are: that no inverse operator has to be determined, the iteration can

be stopped before convergence providing a partially deblurred $\hat{f}(\mathbf{x})$ that does not suffer from the inverse filter's high frequency artifacts, and additional prior information about the form of the solution can easily by used to constrain each $\hat{f}_k(\mathbf{x})$ and improve the final solution. As with the filters described in the previous sections one of the major limitations of this technique is that the quality of $\hat{f}(\mathbf{x})$ is dependent on the accuracy of $\hat{h}(\mathbf{x})$. It is worth noting that the basic iterative scheme described above has been expanded to give new deconvolution techniques, such as the steepest descent and conjugate gradient techniques, where the value of β is changed at each iteration subject to an optimisation criterion in order to provide faster convergence (Biemond *et al.*, 1990).

3.2.3.2 Projection onto Convex Sets

When a reasonable amount of prior information is available about the form of the solution to a deconvolution problem, it is possible to use the method of projection onto convex sets (POCs) to obtain a solution. The method of POCs assumes that each of the different forms of prior information about the form of the solution restricts the solution to lie within a well defined closed convex set that is a subset of the set of all possible images (Youla and Webb, 1982; Youla, 1987; Katsaggelos, 1989; Biemond *et al.*, 1990). A closed convex set is a finite region of Hilbert space where all the images that satisfy a particular property exist (Youla and Webb, 1982; Katsaggelos, 1989). If a particular form of prior knowledge constrains the image estimate to lie within a particular closed convex set C_i, then the required solution to the deconvolution problem is an image that lies within the closed convex set that is the intersection of all the C_i (Youla and Webb, 1982; Youla, 1987), i.e

$$\hat{f}(\mathbf{x}) \in C_0 = \bigcap_{i=1}^{M} C_i,$$

where $\cap$ denotes the set intersection and there are M different *pieces* of prior information available. Some of the types of prior information that constrain $\hat{f}(\mathbf{x})$ to lie within particular closed convex sets are estimates of the image's energy, extent, bandwidth, range of pixel values, and the requirement that $\hat{F}(\mathbf{u})$ must satisfy $G(\mathbf{u}) \approx \hat{H}(\mathbf{x})\hat{F}(\mathbf{u})$ to within the bounds of the contamination

To obtain a solution to a deconvolution problem by the method of POCs, an image that exists in C_0 must be found. To accomplish this an iterative procedure

$$\hat{f}_{k+1}(\mathbf{x}) = (P_1 P_2 \ldots P_M)\hat{f}_k(\mathbf{x}) \tag{3.16}$$

which successively projects the image estimate to within the bounds of each of the convex sets is used (Youla and Webb, 1982; Katsaggelos, 1989). The P_i in Eqn.(3.16) denote the projection operators that project the $\hat{f}_k(\mathbf{x})$ into the region of Hilbert space defined by each of the C_i, and the product of all the P_i defines a nonexpansive mapping (Youla and Webb, 1982; Youla, 1987; Katsaggelos, 1989). The ordering of the P_i is immaterial (Youla and Webb, 1982). The criteria that could be used with POCs to determine when iteration should be terminated are: to terminate the iteration when $\|\hat{f}_{k+1}(\mathbf{x}) - \hat{f}_k(\mathbf{x})\| < \epsilon$, where ϵ is an arbitrary small number (this method produces poor estimates if the algorithm converges slowly)(Trussell, 1983); to evaluate the residual $R = \|G(\mathbf{u}) - H(\mathbf{u})\hat{F}_k(\mathbf{u})\|$ and when $R = \|c(\mathbf{u})\| \pm \epsilon$, where ϵ is an arbitrary number much smaller than $\|c(\mathbf{u})\|$, the iteration is terminated (Trussell, 1983); and to compute the summed distance error $J(\hat{f}_k(\mathbf{x})) = \sum_{i=1}^{M} \|P_i\hat{f}_k(\mathbf{x}) - \hat{f}_k(\mathbf{x})\|$, which describes the Euclidean distance between $\hat{f}_k(\mathbf{x})$ and $\hat{f}_k(\mathbf{x})$ projected into each of the convex sets, and terminate the iteration when $J(\hat{f}_k(\mathbf{x}))$ is less than an arbitrary small number (Levi and Stark, 1987). The first

two criteria could be used with any of the iterative techniques described herein to determine when iteration should be terminated.

The quality of the $\hat{f}(\mathbf{x})$ obtained by POCs is clearly heavily dependent on the accuracy of the prior information and assumptions made about the form of the image (Civanlar and Trussell, 1986). The reason for this is that inaccurate prior information can make C_0 an empty set and less precise prior information could increase the size of C_0 which decreases the likelihood of obtaining an accurate $\hat{f}(\mathbf{x})$. Recently the method of POCs has been extended to a method of projection onto fuzzy sets where the boundaries defining the convex sets are now fuzzy instead of being well defined (Civanlar and Trussell, 1986). With such fuzzy boundaries, the convex sets are easier to define and solutions less sensitive to the accuracy of the assumptions made regarding the form of the solution (Biemond *et al.*, 1990). It is also interesting to note that the constrained Van Cittert iterative deconvolution technique, depending on the type of constraints used, could also be considered to be a solution method using POCs.

3.2.3.3 Kalman Filtering

In the situation where it is possible to assume the following: the original image can be described by a zero-mean wide-band stationary discrete random field, the blurring of the original image was performed by a LTI system, and the contamination can be modelled by a white Gaussian noise source, a solution to the contaminated deconvolution problem can be obtained by using the recursive Kalman filtering technique (Woods and Radewan, 1977; Biemond *et al.*, 1990). The form of the Kalman filter can be derived in the manner now outlined. The above assumptions permit the original 2-D pixellated image to be modelled by a 2-D autoregressive process

$$f[n,m] = \sum_{(i,j)\in\Pi} a[i,j]f[n-i,m-j] + u[n,m], \tag{3.17}$$

which describes the current pixel value of the image in terms of the previous pixel values (i.e. the above model is causal and therefore Π corresponds to the non-symmetric half plane (Woods and Radewan, 1977)), the image model parameters $a[p,q]$, and the noise source $u[n,m]$ that describes the error in approximating the $f[n,m]$ by a linear combination of weighted neighbouring pixels (Woods and Radewan, 1977; Biemond *et al.*, 1990; Katsaggelos, 1991). In a similar manner the blurred and contaminated image can be modelled by the autoregressive model

$$g[n,m] = \sum_{(k,l)\in\Pi} h[k,l]f[n-k,m-l] + c[n,m], \tag{3.18}$$

(cf. Eqn.(2.49)) which is known as the observation model. On solving the above two autoregressive or state equations for the image, the equations

$$\hat{f}_p[n,m] = \sum_{(i,j)\in\Pi} a[i,j]\hat{f}_u[n-i,m-j],$$

and

$$\hat{f}_u[n,m] = \hat{f}_p[n,m] + k[n,m]\left\{ g[n,m] - \sum_{(k,l)\in\Pi} h[k,l]f[n-k,m-l] \right\},$$

defining the form of the Kalman filter, are obtained. In the above Kalman filter equations the $k[n,m]$ are the Kalman filter gain parameters, which can be calculated from the Riccati equation (Katsaggelos, 1991), and the subscripts p and u denote the predicted and updated versions of the image estimate respectively (Woods and Radewan, 1977; Biemond *et al.*, 1990;

Katsaggelos, 1991). In essence the Kalman filter predicts an estimate for the current pixel value by using the autoregressive image model and the image estimate is then refined, using the knowledge of the form of the blurred image and psf, to obtain a refined (i.e. updated) estimate of the pixel value. By moving the filter across the image one pixel at a time the Kalman filter is able to recursively build up a complete estimate of the original image. Before using the Kalman filter to solve the deconvolution problem the form of the Kalman filter must be determined from prior estimates of the image model, psf and noise statistics using techniques such as the technique alluded to in §3.3.5. As a consequence $\hat{f}(\mathbf{x})$ is dependent on the accuracy of these parameters (Biemond *et al.*, 1990; Katsaggelos, 1991). During the process of Kalman filtering an image of finite extent, it is necessary to extend the boundary of the image in a manner which allows the application of the filter. It is usual to extend the boundary of the image using values based on past and present data or using the mean value of the blurred image's pixels (Woods and Ingle, 1981).

3.2.3.4 Maximum Entropy Deconvolution

In many deconvolution problems the prior information used to constrain the form of the solution may be insufficient to provide a single unique solution but instead may result in a single solution being selected from a set of possible solutions. For example, in the situation where the bound set by Eqn.(3.4) is used to constrain the form of the possible solutions to a deconvolution problem, it is very likely that there are a set of solutions that satisfy the criterion specified by Eqn.(3.4). In situations such as this, what is required is a technique that allows the best possible solution to be selected from the set of possible solutions. Techniques, known as regularisation techniques, attempt to find the best solution by establishing which solution minimises some stabilising functional $\Omega\{\hat{f}(\mathbf{x})\}$ (Biemond *et al.*, 1990). As one of the most common types of stabilising function used is the maximum entropy measure, a large family of maximum entropy deconvolution techniques have arisen.

The maximum entropy deconvolution techniques regularise the deconvolution technique by maximising the stabilising functional

$$\Omega\{\hat{f}(\mathbf{x})\} = S = -\sum_{n,m} \hat{f}[n,m] \log \hat{f}[n,m], \tag{3.19}$$

which describes the entropy of the image estimate (Gull and Skilling, 1984; Zhuang *et al.*, 1987). The above formula has its foundations in the field of information theory and describes the amount of *disorder* in a particular solution. The reason why the above maximisation is believed to provide the best solution is that it is unreasonable to expect the error in the pixel values (i.e. the $f[n,m] - \hat{f}[n,m]$) of the best $\hat{f}(\mathbf{x})$ to be confined in an ordered manner to a single pixel or group of pixels; it is far more reasonable to expect the error to be spread evenly over the entire image estimate in a disordered manner (Gull and Skilling, 1984). Consequently, the most disordered image, in the set of images that have the structure necessary to satisfy the constraints specified by the prior knowledge, is considered to be the best $\hat{f}(\mathbf{x})$. Stated another way, maximum entropy techniques provide image estimates that maximally satisfy the prior knowledge whilst being maximally noncommittal about unknown parameters (Gull and Skilling, 1984). It is worth noting that Gull and Skilling (1984) state that the maximum entropy functional is the only function that is known to provide sensible solutions to the regularised deconvolution problem because other functionals tend to distribute the error across the image in a nonsensical manner.

The fact that the maximum entropy techniques tend to spread the error over the entire

image results in image estimates that have a minimum of artifacts, and have an automatic balance between resolution and the adverse effects of high frequency contamination (unlike the inverse and Wiener filters) (Gull and Skilling, 1984). Like the techniques described in the previous sections the accuracy of the solutions provided by the maximum entropy techniques are dependent on the accuracy of the prior information used to constrain the form of $\hat{f}(\mathbf{x})$, since the size of the set of possible $\hat{f}(\mathbf{x})$ is directly related to the accuracy of the prior knowledge. While the relationship between the accuracy of $\hat{f}(\mathbf{x})$ and the accuracy of the prior knowledge may not be as strong as in the case of Wiener and inverse filtering, due to the maximum entropy criterion, the effects of inaccurate prior knowledge are still noticeable. An excellent example of the effects of psf inaccuracies, on the accuracy of solutions provided by maximum entropy deconvolution techniques, is shown in Fig. 1 in Gull and Skilling (1984) where artifacts from incorrectly modelling the psf are clearly evident.

As stated above the maximum entropy technique can be combined with a number of different deconvolution techniques to obtain an improved solution to the deconvolution problem. Consequently, a number of different algorithms exist for finding which image, within the set of images that satisfy the particular constraints, has the maximum entropy. A description of the various techniques used to determine a maximum entropy solution is beyond the scope of this thesis, however, the majority of techniques use an iterative search of the set of possible solutions to find the image that maximises Eqn.(3.19) (Newman and Hildebrandt, 1987; Zhuang *et al.*, 1987). In the situation where the image is blurred by a LTI system, Gull and Skilling (1984) note that the maximum entropy solution is unique, implying that the above iterative search techniques converge on global maxima in this situation.

3.2.4 1-D Zero-Based Consistent Deconvolution

It was indicated in the previous sections that all of the techniques reviewed provide image estimates that have an accuracy that is directly dependent on the accuracy of $\hat{h}(\mathbf{x})$. A deconvolution technique which provides an image estimate with a quality which is less dependent on the quality of $\hat{h}(\mathbf{x})$ is the 1-D zero-based consistent deconvolution technique. The reason is that the latter technique uses $\hat{h}(\mathbf{x})$ for recognition purposes only and does not explicitly use $\hat{h}(\mathbf{x})$ in computing $\hat{f}(\mathbf{x})$. The 1-D consistent deconvolution algorithm was developed by Bates *et al.* (1990) and while this 1-D algorithm has a limited usefulness, the concepts can be applied usefully to deconvolve images of higher dimension (see §5.1).

One-dimensional blurred images are deconvolved by the 1-D consistent deconvolution algorithm by using the principle that it is possible to divide the set of points in Z space where the Z transform of the 1-D image is zero (known as zeros) into two distinct subsets. A complete description of the properties of an image's Z transform zeros is given in Chapter 4. It is also shown in §4.3.1 that one of the zero subsets pertains to the original image and can be used to reconstruct an estimate of the image, whilst the second subset of zeros belongs to the psf. The 1-D consistent deconvolution algorithm uses this property of the zero subsets to make the rather straightforward assumption that the zeros of the Z transform of the psf estimate should be situated in approximately the same region of Z space as the set of blurred image's zeros belonging to the psf. Thus, using the zeros of $\hat{\mathcal{H}}(\zeta)$ it is possible to recognise and eliminate the zeros of $\mathcal{G}(\zeta)$ that belong to $\mathcal{H}(\zeta)$. This process leaves only the zeros of $\mathcal{G}(\zeta)$ belonging to $\mathcal{F}(\zeta)$ and these can be used to obtain $\hat{\mathcal{F}}(\zeta)$. The reason why this algorithm is known as a consistent deconvolution technique is because, in the uncontaminated situation, the algorithm is able to use $\hat{\mathcal{H}}(\zeta)$ to effectively filter $\mathcal{H}(\zeta)$ from the convolution, whereas the techniques described in the previous sections remove $\hat{\mathcal{H}}(\zeta)$ from the convolution. A full account of what

is meant by the term consistent deconvolution is given in §3.4.

In the presence of contamination the zeros of $\mathcal{G}(\zeta)$ are displaced in Z space and, as Bates *et al.* (1990) demonstrate, it becomes more difficult to correctly match the zeros of $\hat{\mathcal{H}}(\zeta)$ to those in the convolution belonging to $\mathcal{H}(\zeta)$. Nonetheless, for reasonable levels of contamination it is still possible to obtain a solution to the 1-D deconvolution problem by using the 1-D consistent deconvolution algorithm (Bates *et al.*, 1990). Bates *et al.* (1990) did not explicitly investigate how sensitive the quality of the $\hat{\mathcal{F}}(\zeta)$ obtained by consistent deconvolution were to the inaccuracy in $\hat{\mathcal{H}}(\zeta)$. In the 1-D case, however, the inaccuracies in the psf estimate would tend to displace the psf zeros to be matched in much the same manner as additive contamination on $g(x)$. For this reason, for reasonable levels of $\hat{\mathcal{H}}(\zeta)$ inaccuracy, the accuracy of $\hat{\mathcal{F}}(\zeta)$ should be reasonably independent of the accuracy of $\hat{\mathcal{H}}(\zeta)$. In fact, the results presented in §5.2 and §5.7 for the 2-D case support this contention.

The concepts used in the 1-D consistent deconvolution algorithm to obtain a solution to the standard deconvolution problem can be extended to higher dimensions. The extension of the algorithm to 2-D is one of the original contributions presented herein (see §5.1). Bates *et al.* (1990) were the first authors to note that, in the contaminated situation, the loci of zeros belonging to the 2-D Z transform of the psf estimate (see §4.2.1) are still recognisably similar to the loci of zeros belonging to the psf in the convolution, since they progress in the same direction. This principle is one of those used in the new 2-D deconvolution algorithm to aid the recognition of the actual psf in the convolution. Bates *et al.* (1990) also noted in their paper that the consistent deconvolution algorithm may be able to be used in conjunction with the Wiener filter to improve the resolution of the latter deconvolution technique and this possibility is discussed in §7.2.

3.3 ESTIMATING THE PSF

All of the deconvolution methods described in §3.2 require an estimate of of the psf in order to be able to obtain a solution to the deconvolution problem. In this section the various techniques that can be used to obtain an estimate of the form of the psf are briefly reviewed. The psf estimation techniques are reviewed here for completeness and to emphasise that the $\hat{h}(\mathbf{x})$ obtained from most of these techniques is likely to be significantly different than $h(\mathbf{x})$; the majority of the techniques assume in their derivation that the blurred image is uncontaminated.

The methods that can be used to recover $\hat{h}(\mathbf{x})$ can be grouped into two categories: spatial or frequency domain techniques. The type of technique used depends on the form of the psf, the nature of the blurring and whether access can be gained to the system used to produce the blurred image. In particular, if the blurring is not isoplanatic spatial domain techniques have to be used since the convolutional model is no longer valid. The psf estimation techniques discussed in this section and their advantages and disadvantages are summarised in Table 3.2.

3.3.1 Optical Transfer Function Zero Recognition

The psfs that have a simple form, such as those describing ideal out-of-focus and linear motion blur, tend to have an an optical transfer function (OTF) that has an oscillatory nature and a characteristic pattern of zero crossings (Gennery, 1973; Bates and McDonnell, 1986; Biemond *et al.*, 1990). The characteristic pattern of these zero crossings may allow a psf estimate to be obtained from $B(\mathbf{u})$ by the method of OTF zero recognition. In §2.4 it was shown that in the

PSF ESTIMATION TECHNIQUES

Technique	Domain	Required PSF Form	Strengths	Limitations
Zero Recognition.	F.	Zero crossings in spectrum, e.g. disc or line psf.	Simple. Cepstrum can highlight zero crossings.	Zeros must be recognisable.
Average Logarithm Spectrum.	F.	$h(\mathbf{x})$ extents $\ll g(\mathbf{x})$ extents.	Reconstructs any isoplanatic psf.	Estimate dependent on $\overline{\log\{\hat{F}_k\}}$.
Edge Recognition.	F.	Any isoplanatic psf.	Reconstructs any isoplanatic psf.	Very sensitive to contamination. Must recognise edges of varying orientation.
Calibration.	F. and I.	Any isoplanatic psf.	Straightforward.	Not always practical.
Autoregressive Modelling.	I.	Any described by a zero mean random field.	Reconstructs non-isoplanatic psfs. Provides image model parameters.	Computationally expensive.

Table 3.2 Key: F. = Fourier and I. = Image.

uncontaminated case the spectrum of the blurred image is simply the product of the spectrum of $f(\mathbf{x})$ and the OTF. A consequence of this relationship is that $B(\mathbf{u})$ is zero whenever the OTF is zero, therefore it may be possible to recognise the distinctive zero patterns of the OTF in $B(\mathbf{u})$ and obtain an estimate of $h(\mathbf{x})$. To be able to recognise the distinctive zero pattern it is necessary for: $F(\mathbf{u})$ to have a bandwidth that extends beyond the region in Fourier space where the recognisable zero pattern of the OTF resides, $F(\mathbf{u})$ not to possess a distinctive zero pattern of its own, $B(\mathbf{u})$ not to be distorted by spectral leakage or aliasing which may mask the zeros of the OTF, and $F(\mathbf{u})$ must not be zero in the neighbourhood of the zeros of the OTF (Gennery, 1973). As an example consider the situation where a 2-D blurred image is blurred by a linear motion blur in the x direction of extent a. In this case the OTF has the form

$$H(u,v) \propto \frac{\sin(\pi a u)}{\pi a u},$$

and $B(u,v)$ is therefore zero at $u = \frac{\pm 1}{a}, \frac{\pm 2}{a}, \frac{\pm 3}{a}, \ldots$. This pattern of zeros should be recognisable, allowing the characteristics of the psf to be estimated, provided that $F(\mathbf{u})$ has the form described above (Gennery, 1973; Andrews and Hunt, 1977; Biemond *et al.*, 1990). It should be clear from the above discussion that if the OTF does not have zero crossings, such as in the case where the psf can be modelled by a Gaussian, this method is of no use (Biemond *et al.*, 1990)

When the blurred image is corrupted by additive contamination the addition of the term $C(\mathbf{u})$ to $B(\mathbf{u})$ tends to perturb the zeros of $B(\mathbf{u})$ making the recognition of the zero pattern of the OTF far more difficult. Thus, the accuracy of the psf estimate provided by this technique is likely to be very sensitive to the level of additive contamination. There are two methods that can be used to accentuate the distinctive zero patterns of the OTF in $G(\mathbf{u})$. The first of these methods can be used in the situation where the blurring of an image is non-isoplanatic (due to the finite depth of field of the imaging system (see §2.3.2)) and/or contains a featured

region of interest surrounded by a smooth region where the blurring is less evident. On extracting the region of interest from the image (by truncation) and computing the spectrum of this extracted image, it is found that the OTF zeros are far more prominent than in the spectrum of the complete image (Biemond *et al.*, 1990). The second method that can be used to emphasise the zeros of the OTF is to compute the cepstrum (the IFT of the logarithm of the spectrum) of the blurred image (Bates and McDonnell, 1986; Biemond *et al.*, 1990). In order to determine the form of a cepstrum of an OTF consider the case where an OTF tends to zero at $u = u_z, 2u_z, 3u_z, \ldots$. As the logarithm of the OTF tends to infinity at the above values of u the cepstrum of the OTF therefore has large spikes at $x = \frac{1}{u_z}, \frac{2}{u_z}, \frac{3}{u_z}, \ldots$. Summation in the cepstral domain is equivalent to convolution in the spatial domain and consequently the OTF zeros manifest themselves as large spikes in the cepstrum of the blurred image, allowing easier recognition of the OTF zeros. In the case of the linear motion blur described previously the cepstrum of $b(x, y)$ has peaks at $x = a, 2a, 3a, \ldots$ which again allows the extent and direction of the blurring function to be recognised (Biemond *et al.*, 1990).

3.3.2 Average Logarithm Spectrum

The average logarithm spectrum method is a technique which allows any type of psf to be estimated from the blurred image when the blurring is isoplanatic. The principle which the average logarithm spectrum method uses to obtain $\hat{h}(\mathbf{x})$ from $g(\mathbf{x})$ is now described. In the uncontaminated situation the power spectral density of the blurred image is given be the expression

$$|B(\mathbf{u})|^2 = |H(\mathbf{u})|^2 |F(\mathbf{u})|^2.$$

If $G(\mathbf{u})$ is now divided up into k subimages, each of which has an extent much larger than the psf estimate, it is found that the power spectral density of these subimages is approximated by the relationship

$$|B_k(\mathbf{u})|^2 \approx |H(\mathbf{u})|^2 |F_k(\mathbf{u})|^2,$$

where the subscript k denotes the kth subimage (Andrews and Hunt, 1977; Biemond *et al.*, 1990). On taking the logarithm of the above expression and averaging the resulting expressions for each of the subimages, the expression

$$\begin{aligned} \overline{\log\{B_k^2\}} &= \frac{1}{K} \sum_{k=1}^{K} \log |B_k(\mathbf{u})|^2 \\ &\approx \frac{2}{K} \sum_{k=1}^{K} \log |F_k(\mathbf{u})| + \log |H(\mathbf{u})|, \end{aligned}$$

is obtained. The term $\overline{\log\{B_k^2\}}$ can be calculated from the blurred image and it has been found that the summation involving the $F_k(\mathbf{u})$ can be approximated reasonably well by a function $\overline{\log\{\hat{F}_k\}}$ which is the average logarithm of the magnitude spectra of a number of arbitrary images (Biemond *et al.*, 1990). Thus, if $\overline{\log\{\hat{F}_k\}}$ is subtracted from $\overline{\log\{B_k^2\}}$ an estimate of the logarithm of the power spectrum of the psf remains, allowing the magnitude response of the blurring function to be estimated. Clearly, the accuracy of the psf estimation scheme is heavily dependent on the level of additive contamination and the form of $\overline{\log\{\hat{F}_k\}}$. This method was first used by (Stockham *et al.*, 1975) who successfully restored old audio recordings of operatic performers.

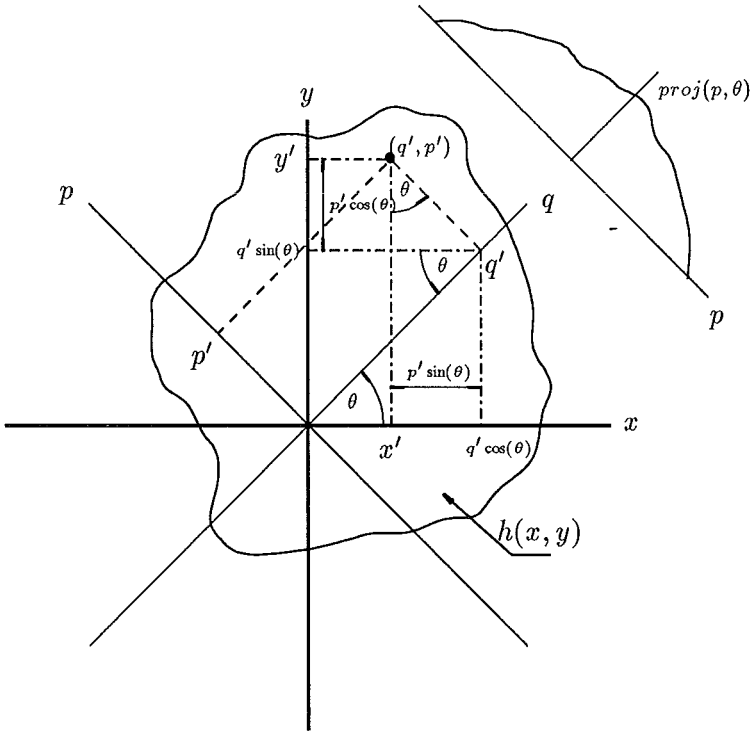

Figure 3.4 The coordinate system (x, y) and its relationship to the coordinate system (p, q) which is used to obtain the projection, $proj(p, \theta)$, of $h(x, y)$.

3.3.3 Edge Recognition

The edge recognition psf estimation technique reconstructs an estimate of the psf from a number of different psf projections. A psf projection is defined by the integral

$$\text{proj}(p, \theta) = \int_{-\infty}^{\infty} h(x, y) dq, \tag{3.20}$$

where (p, q) is a cartesian coordinate system in image space rotated at an angle θ with respect to the coordinate system (x, y), as is shown in Fig. 3.4, and $proj(p, \theta)$ is the integrated psf intensity along a ray parallel to the q axis and intersecting the p axis at a particular value of p. For a complete discussion on projections and reconstruction from projections the reader is referred to Bates and McDonnell (1986).

This psf estimation technique obtains the psf projections required to reconstruct the psf by differentiating the blurred image at the regions in image space where there is a rapid change in the intensity of the image, i.e. at an edge within the image. To understand how the derivatives of the edges provide the projections necessary to reconstruct the psf it is useful to model an edge by the function

$$A \, \text{step}(p - p_0),$$

where the step function is a unit step function, p_0 is the displacement of the step function along the p axis previously defined, and A is the amplitude of the step. In the uncontaminated situation the blurred image in the neighbourhood of the above edge can be described by the expression

$$g_e(x, y) = A \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x', y') \text{step}(p - p_0 - p') dp' dq', \tag{3.21}$$

where $x' = q'\cos(\theta) - p'\sin(\theta)$ and $y' = q'\sin(\theta) + p'\cos(\theta)$. Differentiating Eqn.(3.21) with respect to p gives

$$\left.\frac{\partial g_e(x,y)}{\partial p}\right|_{\text{vicinity of step}} = A\int_{-\infty}^{\infty} h(x'',y'')dq, \tag{3.22}$$

since $\frac{d\text{step}(p)}{dp} = \delta(p)$ (Rosenfeld and Kak, 1982; Bates and McDonnell, 1986) and where $x'' = q'\cos(\theta) - (p - p_0)\sin(\theta)$ and $y'' = q'\sin(\theta) + (p - p_0)\cos(\theta)$.

On comparing Eqn.(3.20) to Eqn.(3.22) it can be seen that, in the uncontaminated situation, the derivative of an edge is indeed equivalent to a projection of the psf. If a number of psf projections are obtained, from edges orientated at various θ values, an estimate of the psf can be obtained by using one of the algorithms that allow an image to be reconstructed from its projections. Like the previous psf estimation techniques, the accuracy of this technique is going to suffer from the initial assumption that the blurred image is uncontaminated. In addition, it has also been found that the differential operator makes the technique very sensitive to contamination such as film grain noise (Rosenfeld and Kak, 1982).

3.3.4 Calibration

Perhaps the most obvious technique that can be used to obtain a psf estimate is the method of calibration. The method of calibration involves using the imaging system, which introduced the blurring, to obtain a reference image of a known object. As the original image is then known *a priori*, the form of the blurring in this reference image can be determined and this in turn provides a psf estimate for the blurred image of interest. This technique can be used in the situation where the form of the psf does not change with time or changes very slowly with time. If the form of the psf changes slowly with time the time interval between obtaining the reference image and the image of interest must be sufficiently short that no appreciable change occurs in the form of the psf.

The form of the psf estimate can be determined from the reference image by using the standard deconvolution algorithms described previously. When using these algorithms to obtain $\hat{h}(\mathbf{x})$ the roles of $\hat{h}(\mathbf{x})$ and $\hat{f}(\mathbf{x})$ are now reversed since $\hat{f}(\mathbf{x})$ is now known *a priori*. The accuracy of the solutions provided by these algorithms is dependent on the level of additive contamination on the reference image (see Table 3.1), consequently the accuracy of $\hat{h}(\mathbf{x})$ is again dependent on the level of the additive contamination. The major limitation of this technique is that the imaging system used to obtain the original blurred image is not always available for the purpose of obtaining a reference image.

3.3.5 Identification by Autoregressive Modelling

The final method presented here for obtaining $\hat{h}(\mathbf{x})$ is the autoregressive modelling technique (Tekalp *et al.*, 1986; Biemond *et al.*, 1988; Biemond *et al.*, 1990). This method is particularly useful as it is able to obtain a model for non-isoplanatic blurring functions and can provide the image model parameters required by the Kalman filter (see §3.2.3.3) (Biemond *et al.*, 1988). The technique uses the autoregressive model for the original image in Eqn.(3.17) and the autoregressive model for the blurred image in Eqn.(3.18), the latter with non-causal support, to obtain an expression for the psf model parameters. The model for the psf is obtained by eliminating the unknown original image from Eqn.(3.17) and Eqn.(3.18). This elimination is

performed by substituting Eqn.(3.17) into Eqn.(3.18) to give

$$g[n,m] = \sum_{(i,j)\in\Pi} a[i,j] \sum_{(k,l)\in\Pi} h[k,l]f[n-i-k, m-j-l] + \sum_{(k,l)\in\Pi+} h[k,l]u[n-k, m-l] + c[n,m], \quad (3.23)$$

and then substituting Eqn.(3.18) into the second summation in the first term of Eqn.(3.23) to give

$$g[n,m] = \sum_{(i,j)\in\Pi} a[i,j]g[n-i, m-j] - \sum_{(i,j)\in\Pi} a[i,j]c[n-i, m-j] + \sum_{(k,l)\in\Pi+} h[k,l]u[n-k, m-l] + c[n,m], \quad (3.24)$$

where $\Pi+$ denotes that the summation occurs over the entire subimage of interest and not just the non-symmetric half plane, i.e. the model is non-causal (Tekalp *et al.*, 1986).

In the situation where the additive contamination on the blurred image is negligible Eqn.(3.24) reduces to the expression

$$g[n,m] = \sum_{(i,j)\in\Pi} a[i,j]g[n-i, m-j] \sum_{(k,l)\in\Pi+} h[k,l]u[n-k, m-l],$$

which is an autoregressive moving average model for the blurred image (Tekalp *et al.*, 1986; Biemond *et al.*, 1990). A discussion of the techniques which are used to obtain the image coefficients $a[p,q]$ and blur coefficients from the ARMA model is beyond the scope of this thesis and the interested reader is referred to Tekalp *et al.* (1986) and Biemond *et al.* (1990). The assumption regarding the magnitude of the contamination used to obtain the ARMA model is not likely to be valid in many imaging situations and, as Tekalp *et al.* (1986) note, the psf estimate and image model parameters found by this technique are perturbed by this contamination.

3.4 CONSISTENT DECONVOLUTION

A deconvolution technique is said to be consistent if it removes the actual blurring from a blurred image and not simply an estimate of the blurring. None of the 2-D deconvolution techniques described in §3.2 are consistent. In this section a more formal description of the conditions that a deconvolution algorithm must satisfy in order to be classed as a consistent deconvolution algorithm is given. This discussion should in turn explain why the term consistent is used to describe those techniques that can use $\hat{h}(\mathbf{x})$ to filter $h(\mathbf{x})$ from $b(\mathbf{x})$.

The consistent deconvolution principle was initially defined by Bates *et al.* (1976) and the description of the principle presented here follows their argument. Consider two 1-D pixellated images $f(x)$ and $h(x)$ which have extents N and M respectively. The discrete convolution of these images is given by the expression

$$b[r] = \sum_{n=0}^{M-1} h[n]f[r-n] \quad r = 0, 1, \ldots, N+M-2,$$

where the $b[r]$ are the pixel values of the resulting image $b(x)$. On expanding the above convolution expression a system of equations of the form

$$b[0] = h[0]f[0]$$

$$\begin{aligned} b[1] &= h[0]f[1] + h[1]f[0] \\ &\vdots = \vdots \\ b[N-1] &= h[0]f[N-1] + h[1]f[N-2] + \ldots + h[M-1]f[N-M] \\ &\vdots = \vdots \\ b[N+M-2] &= h[M-1]f[N-1], \end{aligned}$$

is obtained. If it is now assumed that $b(x)$ and an estimate of $h(x)$ are known, as in the standard deconvolution case, the above system of equations becomes an overdetermined system of $N+M-1$ equations in N unknowns. Solving N of these $N+M-1$ equations for $f(x)$, in the situation where $\hat{h}(x) \neq h(x)$, gives a solution which when substituted into the remaining $M-1$ equations provides values for the $b[r]$ that do not in general agree with the $b[r]$ known *a priori*. The $b[r]$ computed from $\hat{h}(x)$ and $\hat{f}(x)$ are not consistent with the observed $b[r]$ and therefore the solution obtained is described as an inconsistent solution (Bates *et al.*, 1976).

To see how this argument relates to the standard deconvolution problem consider the case where the 2-D algorithms described in §3.2 are used to provide a solution to the uncontaminated standard deconvolution problem when $\hat{h}(\mathbf{x}) \neq h(\mathbf{x})$. If the $\hat{f}(\mathbf{x})$ provided by these techniques could be combined with $h(\mathbf{x})$ to produce equations for the $b[\mathbf{r}]$ of the same form as above, it would be found that the majority of the $b[\mathbf{r}]$ would be inconsistent with the known $b[\mathbf{r}]$ since $\hat{f}(\mathbf{x}) \neq f(\mathbf{x})$. These standard deconvolution algorithms therefore do not provide consistent solutions to the deconvolution problem and cannot be classified as consistent deconvolution techniques. Conversely, the 1-D consistent deconvolution algorithm is a consistent technique because, as explained in §3.2, it uses $\hat{h}(x)$ to recognise and remove the actual psf from the blurred image. The $\hat{f}(x)$ obtained from this technique therefore agrees with $f(x)$ to within the numerical accuracy of the computer used and if substituted into the overdetermined system of equations would provide $b[r]$ that agree with the observed $b[r]$.

In §14 of Bates *et al.* (1976) it is demonstrated that the consistent deconvolution principle can be shown in a more transparent manner, than that used above, by considering the zeros of the Z transforms of the images. As is explained fully in §4.3 and alluded to in §3.2.4, the zeros of an uncontaminated blurred image's Z transform consist of two separate sets of zeros which belong to the Z transform of the original image and psf respectively. Consequently, unless the union of the sets of Z transform zeros belonging to the $\hat{f}(\mathbf{x})$ and $\hat{h}(\mathbf{x})$ is approximately equal to the set of Z transform zeros belonging to $b(\mathbf{x})$, $\hat{h}(\mathbf{x})$ and/or $\hat{f}(\mathbf{x})$ are not the images used to produce the blurred image, i.e. the Z transform zeros of $\hat{h}(\mathbf{x})$ and/or $\hat{f}(\mathbf{x})$ are inconsistent with those of $b(\mathbf{x})$. The condition that the union of the two sets of zeros belonging to the estimated component images must be approximately equal to the set belonging to $b(\mathbf{x})$, if the component images are consistent with the blurred image, is known as the consistency condition (Bates *et al.*, 1976; Bates and McDonnell, 1986). The 1-D consistent deconvolution algorithm has the consistency condition inherently built in and this is the reason why this technique provides consistent solutions to the uncontaminated deconvolution problems.

3.5 PHASE RETRIEVAL TECHNIQUES

The Fourier phase retrieval problem, as was stated in §3.1, can be posed as: given $|F(\mathbf{u})|^2$ recover an estimate of $f(\mathbf{x})$. The problem of recovering an image from its Fourier transform magnitude arises in many practical imaging situations, such as the areas of antenna aperture imaging and crystallography, as was alluded to in Chapter 1 and §2.3.1.2. Other imaging

situations where the phase retrieval problem occurs are the areas of speckle imaging, ultrasonic diffraction tomography, electron microscopy and wavefront sensing (Fienup, 1982; Bates and Mnyama, 1986). The complete Fourier transform of $f(\mathbf{x})$ (i.e. including the phase) cannot be measured in the above imaging areas for one of or a combination of the following reasons, as outlined by Lane (1988):

1. The receivers which can be used to measure the phase of high frequency electromagnetic fields are expensive or difficult to obtain.

2. The propagation medium through which the electromagnetic waves must travel, before its phase is measured, may distort the phase of the wave, e.g. atmospheric distortion.

3. The instrument used to measure the phase may have inherent inaccuracies, such as misaligned panels in the antenna used to measure $F(\mathbf{u})$, which distort the measured phase.

4. A phase reference may be difficult to maintain when measurements are made at geographic locations that are widely separated.

In this section the various iterative algorithms which can be used to solve the phase retrieval problem are described and the methods which can obtain a direct solution are briefly mentioned. In addition, the form and unique nature of the solutions provided by these algorithms is also described.

3.5.1 Uniqueness of Solution

3.5.1.1 Image Form

When the phase of an image cannot be measured, some information about the image is irretrievably lost (Bates and Mnyama, 1986). This lost information describes the scale, orientation, and position of the original image in image space. To understand that it is only ever possible to recover what is known as the *image form* of the original image it is necessary to consider the effect of two sets of transformations on the original image $f(\mathbf{x})$. Consider the image defined by

$$\tilde{f}_1(\mathbf{x}) = f(\mathbf{x} - \mathbf{x}_0)\exp^{i\varphi_0}, \tag{3.25}$$

and the image

$$\tilde{f}_2(\mathbf{x}) = f^*(-\mathbf{x}), \tag{3.26}$$

where $\mathbf{x}_0$ is a position in space and φ_0 is a real constant. The first equation, Eqn.(3.25), describes a version of the original image that has been translated in image space and multiplied by a complex scale factor, whilst Eqn.(3.26) describes a version of the original image that has been reflected in the origin of the coordinate axis and then conjugated, i.e. it is the conjugate reflection. The images defined by Eqns. (3.25) and (3.26) are instantly recognisable as being simple transformations of $f(\mathbf{x})$ as both transformed images have the same structure as $f(\mathbf{x})$. For this reason, images such as these are said to have the same image form as $f(\mathbf{x})$. It should be noted that the transformations in Eqns. (3.25) and (3.26) can be combined in any manner to produce other images which are image forms of $f(\mathbf{x})$.

If the Fourier transform magnitudes of Eqns. (3.25) and (3.26) are computed it is found for real $\mathbf{u}$ (in practice $|F(\mathbf{u})|^2$ is measured where $\mathbf{u}$ is real) that

$$\begin{vmatrix}\tilde{F}_1(\mathbf{u})\\ \tilde{F}_2(\mathbf{u})\end{vmatrix} = \begin{vmatrix}F(\mathbf{u})\exp^{i(\varphi_0+2\pi\mathbf{x}_0\cdot\mathbf{u})}\\ |F^*(\mathbf{u}^*)|\end{vmatrix} = \begin{matrix}|F(\mathbf{u})|\text{, and}\\ |F(\mathbf{u})|\text{.}\end{matrix}$$

It can therefore be seen that all the image forms of $f(\mathbf{x})$ possess the same Fourier magnitude as $f(\mathbf{x})$. Thus, if the phase of $F(\mathbf{u})$ cannot be measured, it is not possible to determine to which of the image forms of $f(\mathbf{x})$ the measured $|F(\mathbf{u})|$ corresponds (Bates and Mnyama, 1986; Bates and McDonnell, 1986). It is therefore necessary to redefine the phase retrieval problem as: given $|F(\mathbf{u})|^2$ recover an estimate of the image form of $f(\mathbf{x})$.

3.5.1.2 Uniqueness

In order to determine whether it is possible to obtain a unique solution to the phase retrieval problem it is first necessary to determine how $|F(\mathbf{u})|^2$ must be sampled if it is to characterise $f(\mathbf{x})$. In §2.4 it was shown that the power spectrum of an image is equivalent to the Fourier transform of the auto-correlation of the image. The integral which defines the auto-correlation of the original image (Eqn.(2.55)) demonstrates that if the image to be considered has compact support $L_{\mathbf{x}}$ then the auto-correlation function has an extent $2L_{\mathbf{x}}$ (Note that if the extent of a pixellated image in one direction is N, the auto-correlation has extent $2N-1$). Consequently, if an image estimate is to be computed via its auto-correlation and power spectrum, it is necessary to sample $|F(\mathbf{u})|^2$ at intervals of no greater than $\frac{1}{2L_{\mathbf{x}}}$ to avoid time domain aliasing, according to the sampling theorem (Bates and Mnyama, 1986). The above result also demonstrates that for compact $f(\mathbf{x})$, the extent of the auto-correlation obtained by $\mathsf{F}^{-1}\{|F(\mathbf{u})|^2\}$ can be used to determine the extent of $f(\mathbf{x})$, since if the reconstructed auto-correlation has extents $2L_{\mathbf{x}}$, the image has extents $L_{\mathbf{x}}$. It should be noted that it is not necessary to know the extents of the image before determining an appropriate sampling interval, because a sampling interval satisfying the above requirements can be deduced by sampling $|F(\mathbf{u})|^2$ at smaller and smaller intervals until $\mathsf{F}^{-1}\{|F(\mathbf{u})|^2\}$ has values that approach zero at the edges of its support (Bates and McDonnell, 1986).

In the first instance it would appear that there are an infinity of images of the form

$$\hat{f}(\mathbf{x}) = \mathsf{F}^{-1}\{|F(\mathbf{u})|\exp^{i\varphi(\mathbf{u})}\},$$

where $\varphi(\mathbf{u})$ is a random phase function, which could be solutions to the phase retrieval problem since they all possess the required power spectrum. The fact that the auto-correlation can provide the extent of $f(\mathbf{x})$ suggests, however, that the majority of the $\hat{f}(\mathbf{x})$ can be eliminated as possible solutions because only a few $\hat{f}(\mathbf{x})$ are likely to have the required extent. An outline is now given of the theory which demonstrates that in two or more dimensions all the $\hat{f}(\mathbf{x})$ which satisfy the image support constraints, imposed by the auto-correlation, are image forms of $f(\mathbf{x})$ and it is therefore possible to obtain a unique solution to the phase retrieval problem as posed in §3.5.1.1. A more detailed and straightforward description of the unique solution of the phase retrieval problem is given in §4.3.1 in terms of what is known as the zero sheet of the auto-correlation. In §4.3.1 the uniqueness question is addressed in terms of the blind deconvolution problem of which the phase retrieval problem is a special case.

To demonstrate that in two or more dimensions the phase retrieval problem has a unique solution, it is necessary to consider the Z transform of the auto-correlation of the image,

$\mathcal{Z}_{ff}(\zeta,\gamma)$ and its properties. In this discussion only the two dimensional case is considered, however the results discussed apply equally to images of higher dimension. In §2.7 it was shown that $\mathcal{Z}_{ff}(\zeta,\gamma)$ was simply the product of the Z transforms of the original image and its conjugate reflection. Solving the phase retrieval problem can therefore be considered to be equivalent to the problem of factorising $\mathcal{Z}_{ff}(\zeta,\gamma)$ into its two component polynomials $\mathcal{F}(\zeta,\gamma)$ and $\mathcal{F}^*(\frac{1}{\zeta^*},\frac{1}{\gamma^*})$. These two component polynomials automatically satisfy the extent constraints imposed by the auto-correlation. It has been shown by Bruck and Sodin (1979) that in $\geq$ 2-D it is almost always impossible to factorise $\mathcal{Z}_{ff}(\zeta,\gamma)$ beyond its component polynomials $\mathcal{F}(\zeta,\gamma)$ and $\mathcal{F}^*(\frac{1}{\zeta^*},\frac{1}{\gamma^*})\zeta^{N-1}\gamma^{M-1}$ (also see 4.3.1). As a consequence, $\mathcal{Z}_{ff}(\zeta,\gamma)$ can only be factorised into the two polynomials which describe an image form of $f(x,y)$ and it conjugate reflection, which in turn indicates that the phase retrieval problem almost always has a unique solution in two or more dimensions. The uniqueness property is not absolute; there are a number of contrived situations, such as the case where $f(x,y)$ is itself an auto-correlation of another image or possesses certain symmetries unlikely to arise in practice, where it is possible to factorise $\mathcal{Z}_{ff}(\zeta,\gamma)$ and obtain many solutions to the problem (Bates and Mnyama, 1986). It is important to note that in 1-D $\mathcal{F}(\zeta)$ and $\mathcal{F}^*(\frac{1}{\zeta^*})\zeta^{N-1}$ can be factorised into many factors and for this reason the solutions obtained to the 1-D phase retrieval problem are not unique.

3.5.2 Phase Retrieval Algorithms

3.5.2.1 General Iterative Loop

The foundation for many of the iterative phase retrieval algorithms is the general iterative loop shown in Fig. 3.5 (Fienup, 1979; Fienup, 1982; Lane, 1988; Davey, 1989; McCormack *et al.*, 1990), the basic idea for which was originally proposed by Gerchberg and Saxton (1972). The Gerchberg-Saxton algorithm is described in detail in §3.5.2.2. The loop basically alternates between image and Fourier space, via Fourier transformation, applying constraints in each domain in an attempt to urge the image estimate towards an image that accurately represents a solution to the particular phase retrieval problem. The iterative loop is a constrained optimisation technique as it constrains the solution at each iteration in an attempt to minimise the error between the image estimate and the original image. It is a valuable exercise to compare the iterative loop to that of the method of POCs described in §3.2.3.2, as it can be seen that the two techniques are very similar. However, the fact that the iterative phase retrieval techniques use magnitude constraints, which do not form convex sets, precludes the algorithms based on the iterative loop from being classified as a method of POCs (Sezan and Stark, 1982; Fienup and Wackerman, 1986).

To start the iterative algorithm, an initial starting estimate of the image, $\hat{f}_0(\mathbf{x})$, is passed to the Fourier transformation process in the algorithm. It is noted by Bates and Mnyama (1986) and Lane (1988) that the best results are obtained from $\hat{f}_0(\mathbf{x})$ which result in an initial random phase since the structure of some non-random $\hat{f}_0(\mathbf{x})$ (e.g. those possessing particular forms of symmetry) can retard the convergence of the algorithm. The $\tilde{F}_0(\mathbf{u})$ formed from the initial estimate is then forced to obey the Fourier space constraints which results in the spectrum $\hat{F}_0(\mathbf{u})$. The constraints usually applied in Fourier space are the known Fourier magnitude and any partial information known about the phase of the actual image. Partial phase information is usually available in those situations where the propagation medium distorts the passage of the waveform such as speckle imaging and ultrasonic diffraction tomography (Bates and Mnyama, 1986). After inverse Fourier transforming $\hat{F}_0(\mathbf{u})$ the image space constraints are applied to the image estimate to obtain the updated image estimate. This updated image

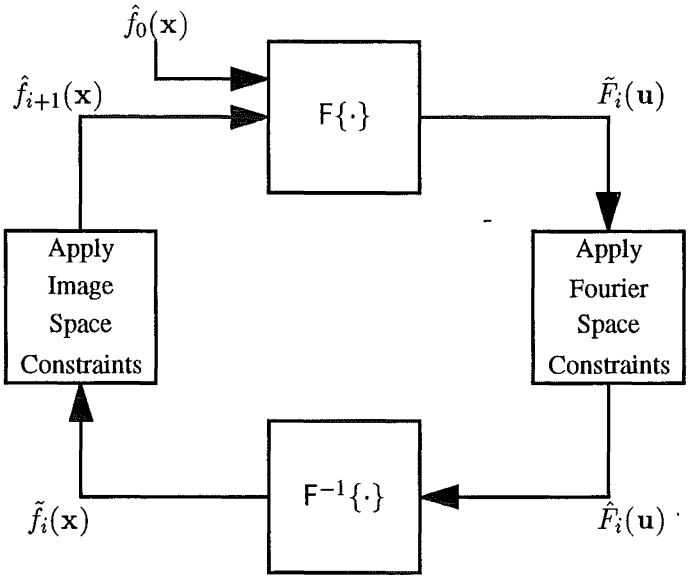

Figure 3.5 The general iterative phase retrieval loop.

estimate is then fed back into the beginning of the loop and the process is continued until a satisfactory estimate of the image is obtained. The most common constraints used in image space are the support constraint, which specifies that the image estimate must be compact and have the extent specified by the auto-correlation, and the positivity constraint, that specifies that the image estimate must be non-negative and real.

As explained earlier the general iterative loop is a constrained optimisation technique. The optimisation that would be ideal is that of minimising the true error, which is the difference between $\hat{f}_{i+1}(\mathbf{x})$ and $f(\mathbf{x})$, or equivalently that of maximising the quality factor, Q, defined in §2.1. In practical phase retrieval situations it is not possible to compute Q or the true error, ϵ_t, as $f(\mathbf{x})$ is unknown. An optimisation goal that can be used in practice is to minimise the violation error defined by the expression

$$\epsilon_v = \frac{\int_{x,y\in\Upsilon} |\tilde{f}_i(\mathbf{x})|^2 d\mathbf{x}}{\int_{-\infty}^{\infty} |\tilde{f}_i(\mathbf{x})|^2 d\mathbf{x}}, \tag{3.27}$$

where Υ denotes the set of points in image space where $\tilde{f}_i(\mathbf{x})$ violates the image space constraints. Another error which can be calculated in practice is the Fourier space error

$$\epsilon_F = \frac{\int_{-\infty}^{\infty} \left(|F(\mathbf{u})| - |\tilde{F}_i(\mathbf{u}|\right)^2 d\mathbf{u}}{\int_{-\infty}^{\infty} |F(\mathbf{u})|^2 d\mathbf{x}}, \tag{3.28}$$

which describes the amount by which $\tilde{F}_i(\mathbf{u})$ violates the Fourier magnitude constraint. The results from the new phase retrieval algorithm in §6.2.1 are presented in terms of ϵ_t, since the results relate to simulations and it is therefore possible to calculate ϵ_t. The matter of the convergence of the iterative phase retrieval algorithm is discussed in the following sections dealing with specific algorithms, since the various methods of enforcing the constraints dictate how the algorithm converges.

3.5.2.2 Gerchberg-Saxton Algorithm

The Gerchberg-Saxton iterative phase retrieval algorithm was originally developed to solve the phase retrieval problem in electron microscopy. The algorithm uses the general iterative loop outlined in §3.5.2.1 and imposes the following image and Fourier domain constraints in an attempt to recover the unknown phase. In Fourier space the magnitude of the estimated spectrum is constrained in a manner that makes it equivalent to the measured Fourier magnitude, and in image space prior knowledge of the magnitude of $f(\mathbf{x})$ is used to constrain the magnitude of the image estimate (Gerchberg and Saxton, 1972; Fienup, 1982; Bates and Mnyama, 1986). Thus, one iteration of the Gerchberg-Saxton algorithm can be described by the set of equations

$$\begin{aligned} \tilde{F}_i(\mathbf{u}) &= \mathsf{F}\{\hat{f}(\mathbf{x})\} \\ \hat{F}_i(\mathbf{u}) &= |F(\mathbf{u})|\exp^{iPh(\tilde{F}_i(\mathbf{u}))} \\ \tilde{f}_i(\mathbf{x}) &= \mathsf{F}^{-1}\{\hat{F}_i(\mathbf{u})\} \\ \hat{f}_{i+1}(\mathbf{x}) &= |f(\mathbf{x})|\exp^{iPh(\tilde{f}_i(\mathbf{x}))}. \end{aligned} \tag{3.29}$$

It is worth noting that this algorithm can be applied to the antenna phase retrieval problem if the design aperture magnitude is used to provide the necessary image space constraints (Gardenier, 1990).

The iteration described by the above set of equations cannot provide estimates that diverge from the actual solution, but this does not guarantee convergence since the algorithm can lock onto a particular estimate which changes very slowly or not at all with each iteration as first shown by Gerchberg and Saxton (1972). The algorithm is then said to have stagnated.

3.5.2.3 Fienup's Algorithms

Several algorithms have been developed by Fienup (1982) to solve the Fourier phase problem in the situation where no prior information is available about $|f(\mathbf{x})|$ but the image space constraints specified in §3.5.2.1 are known. The algorithms of Fienup replace the last equation in the set of equations defining the Gerchberg-Saxton algorithm (Eqn.(3.29)) by an equation that imposes the image space constraints. Fienup's error reduction and input output algorithms use two different approaches to apply the image space constraints. These two techniques are now discussed with reference to the methods used to apply the image space constraints. The other methods by Fienup, which are not described fully here, are gradient search methods that attempt to minimise ϵ_F at each iteration. These gradient search methods, such as the steepest descent method, minimise ϵ_F as a function of the N^2 pixel values representing the image estimate subject to the image space constraints.

In the *error reduction* algorithm the image space constraints are imposed by the equation

$$\hat{f}_{i+1}(\mathbf{x}) = \begin{cases} \tilde{f}_i(\mathbf{x}) & (\mathbf{x}) \notin \Upsilon \\ 0 & (\mathbf{x}) \in \Upsilon, \end{cases} \tag{3.30}$$

where Υ is defined as it was for Eqn.(3.27), which replaces Eqn.(3.29) in the Gerchberg-Saxton algorithm. Like the Gerchberg-Saxton algorithm, Fienup has shown that the error in the image estimate must decrease or remain constant with each iteration (Fienup, 1982). Thus, the error reduction (ER) algorithm can stagnate but not diverge. It has also been noted by Fienup that ϵ_t initially decreases quite quickly, but the rate of convergence then tends to slow and can reach

a number of plateaus before finally converging. The rate of convergence has been found to be dependent on the type of constraints imposed (Fienup, 1982). It has been shown that the ER algorithm is in fact a *steepest descent* method and is a straightforward implementation for such a method (Fienup, 1982).

In an effort to speed up the rate of convergence of the ER algorithm and overcome the stagnation problems, Fienup developed a series of algorithms known as *input output* algorithms, so called because the first three steps of the iterative loop can be considered to be a nonlinear system with an input $\hat{f}_i(\mathbf{x})$ and an output $\tilde{f}_i(\mathbf{x})$ (see Fig.3.6). If small changes are made to the input it has been found that a correspondingly small change occurs in the same direction in the output. Thus, to solve the phase retrieval problem the image estimate is driven towards $f(\mathbf{x})$ by making appropriate small changes to the input which tend to coax the output in the required direction (Fienup, 1982).

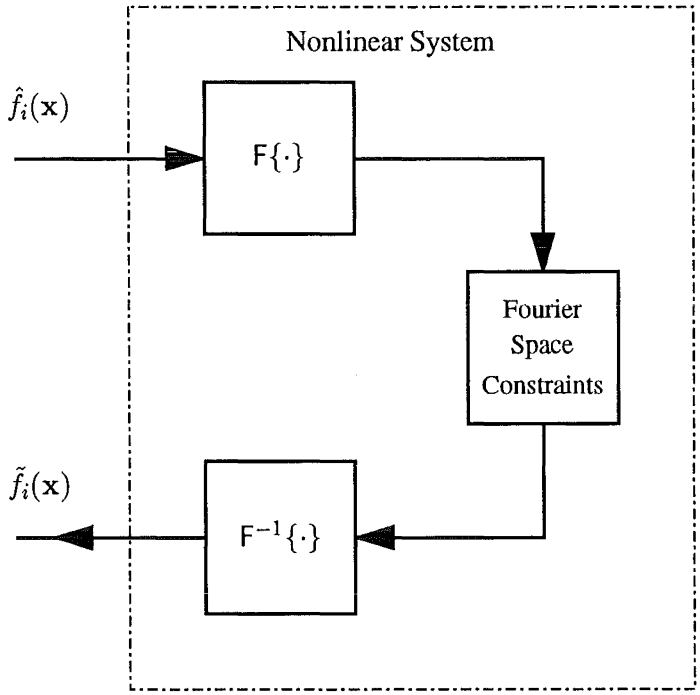

Figure 3.6 A schematic showing the structure of the nonlinear component within the input output family of phase retrieval algorithms.

A reasonably intuitive means of implementing the above principle is to replace Eqn.(3.29) by the equation

$$\hat{f}_{i+1}(\mathbf{x}) = \begin{cases} \hat{f}_i(\mathbf{x}) & (\mathbf{x}) \notin \Upsilon \\ \hat{f}_i(\mathbf{x}) - \beta \tilde{f}_i(\mathbf{x}) & (\mathbf{x}) \in \Upsilon. \end{cases} \tag{3.31}$$

This equation in conjunction with the remaining equations in the Gerchberg-Saxton algorithm define the basic input output algorithm. In Eqn.(3.31) β is a feedback parameter which is a real constant that usually has a value between 0 and 1. It can be seen that the basic input output algorithm implements the above principle by perturbing those pixels that violate the constraints in a direction which should reduce the violation and coax the image estimate towards $f(\mathbf{x})$. The nonlinear terms in the output from the nonlinear system, for a particular input, cannot be predicted exactly and consequently a thorough analysis of the family of input output algorithms presently eludes researchers (Fienup, 1982; Davey, 1989).

Another method of implementing the image space constraints in an input output algorithm is to use the equation

$$\hat{f}_{i+1}(\mathbf{x}) = \begin{cases} \tilde{f}_i(\mathbf{x}) & (\mathbf{x}) \notin \Upsilon \\ \hat{f}_i(\mathbf{x}) - \beta \tilde{f}_i(\mathbf{x}) & (\mathbf{x}) \in \Upsilon. \end{cases}, \tag{3.32}$$

which together with the remaining equations in the Gerchberg-Saxton algorithm define the *hybrid input output* (HIO) algorithm. With the HIO algorithm there is no guarantee that ϵ_t decreases with each iteration and as noted by Davey (1989) ϵ_t tends to fluctuate as the algorithm converges to a solution. The HIO algorithm is, however, reported to be the most effective of all the iterative phase retrieval algorithms as it usually converges to a solution much faster than the other techniques (Fienup, 1982; Fienup and Wackerman, 1986; Lane, 1988).

It was originally believed that the HIO algorithm would not stagnate because as long as the image estimate violates the constraints, the input to the non-linear system should change in a manner that attempts to reduce this violation. In practice the HIO algorithm does stagnate, although not as regularly or as severely as the ER algorithm (Davey, 1989). The types of image estimate that induce stagnation in the HIO algorithm and the techniques that can be used to overcome the stagnation problem are described in Fienup and Wackerman (1986). It is worth noting that one of the instances where stagnation arises is, the twin image problem, where the algorithm simultaneously attempts to reconstruct the image and its conjugate reflection.

3.5.2.4 Other Phase Retrieval Algorithms

There are a number of algorithms, not based on the general iterative loop, which can be used to solve the phase retrieval problem. The techniques most relevant to the work presented in this thesis are those direct techniques that attempt to factorise the Z transform of the auto-correlation, either directly or via the zeros, to obtain an image estimate. Techniques of this type have been developed by Napier and Bates (1974), Nieto Vesperinas and Dainty (1986), Lane *et al.* (1987), and Izraelevitz and Lim (1987). The principle of the method of Lane *et al* and Izraelevitz and Lim are described in §4.3.2 and it is shown that these techniques have the limitation that they cannot be used when the power spectrum has been contaminated. In the contaminated situation the most promising application of the zero based factorisation methods appears to be in assisting the iterative techniques of Fienup in obtaining a solution. There are currently two algorithms which can assist Fienup's algorithms in this way. The first of these is the algorithm developed by Parker and Bones (1992) and this algorithm is described in §6.1.2 along with the second algorithm recently developed by the author.

There are a number of other phase retrieval techniques which are not of direct relevance to the work presented here. These techniques are only briefly described and the interested reader is referred to the references cited for a more detailed description of the algorithms. The maximum entropy phase retrieval algorithm is one of these techniques (Gull and Skilling, 1984). Like the maximum entropy deconvolution algorithm in §3.2.3.4 the phase retrieval technique, in an iterative manner, attempts to find a solution which satisfies the imposed constraints and maximises the entropy function in Eqn.(3.19). The simulated annealing phase retrieval algorithm (Nieto Vesperinas and Mendez, 1986; Nieto Vesperinas *et al.*, 1988) attempts to find the global minimum of the function that describes the error between the measured auto-correlation and the auto-correlation computed from all the possible image estimates. The algorithm finds the minimum of the above error function by perturbing the image estimate to see whether the error function increases or decreases, and if the error function decreases

the perturbation is included within the new image estimate. The final technique to be alluded to here is the direct holographic phase retrieval technique which can be used for retrieving the phase of an antenna's far field radiation pattern (Bates and Napier, 1971; Napier and Bates, 1973). To use this technique the superimposed power spectra of the antenna of interest and a much smaller 'point source' antenna close by must be measured. It transpires that the inverse Fourier transform of the power spectrum contains three separate images which ideally do not overlap in Fourier space. One of these images, which can be recognised by inspection, represents the complex conjugate of the aperture distribution of interest.

CHAPTER 4

ZERO SHEETS

In this chapter the concept of the zero sheet, which is a mathematical representation of an image, is introduced. The zero sheet of a compact image has many useful properties. These properties are discussed in this chapter and were used to advantage when developing the new zero-based deconvolution and phase retrieval algorithms described in §5.1 and 6.1.2.

In the first section of this chapter, §4.1, the mathematical abstraction which is the zero sheet of an image is formally introduced. Many of the basic properties of the zero sheet, including its form and unique relationship to the image from which it was derived, are also discussed in this section. In this first section the discussion is primarily limited to the properties of the zero sheets of a group of images known as irreducible images. An irreducible image is an image which cannot be formed by convolving together any number of component images; except when one of the component images is the image itself and the remaining images are 2-D delta functions (Bates *et al.*, 1990), i.e an irreducible image $a(\mathbf{x})$ cannot be expressed in the form

$$a(\mathbf{x}) = u_1(\mathbf{x}) \odot u_2(\mathbf{x}) \odot \cdots \odot u_N(\mathbf{x}), \quad \text{where } u_i(\mathbf{x}) \neq \delta(\mathbf{x}).$$

Conversely, a reducible image is any image which can be expressed in the form of the convolution shown above, e.g. the blurred image $b(x, y)$ in §2.3.2 is a reducible image.

The dimensionality of the space in which the zero sheets of 2-D images exist make it impossible to display the entire zero sheet of a 2-D image on 2-D media. For this reason techniques have been developed, primarily by others (Lane and Bates, 1987; Bates *et al.*, 1990; Scrase *et al.*, 1992), to allow a 2-D representation of sections of the zero sheet to be formed. These techniques, which allow the zero sheet to be displayed and visualised, are discussed in §4.2. The symmetries possessed by the zero sheets of various forms of irreducible image are also discussed in this section since the display techniques introduced in this section elucidate these properties. In §4.3 the form and properties of the zero sheets of images that have been blurred in the manner described in §2.3.2 or that represent the power spectrum of an image (see §2.4 and §3.5) are introduced. It is discussed in this section how the properties of these two specific forms of reducible image provide a mechanism for deblurring blurred images or reconstructing images from their power spectra. The effect of noise on both the zero sheets of reducible and irreducible images is discussed in §4.4. This section describes: how the noise destroys the useful properties of the zero sheets of reducible images introduced in §4.3, how band-limited noise perturbs different sections of the zero sheet differently, and how the magnitude of the noise in Fourier space relates to its magnitude in Z space. Finally, §4.5 introduces two different techniques, one based around solving an overdetermined system of linear equations and one based around the DFT, which can be used to reconstruct an image from samples of its zero sheet. A comparison of the computational requirements, accuracy, and usefulness of both techniques in various situations is given.

4.1 THE ZERO SHEET OF AN IMAGE

4.1.1 The Zero Sheet and its Properties

The zero sheet of an image is the set of points in Fourier space or Z space where the Fourier transform or Z transform of the image is zero respectively. As this thesis is for the most part concerned with compact 2-D pixellated images it is natural for the following discussion on zero sheets and their properties to be centred around the set of zeros that belong to the image's Z transform.

In Z space the zero sheet of the 2-D irreducible compact image $a(x, y)$ is defined to be the surface, comprising the infinite set of points $\{(\zeta_1, \gamma_1), (\zeta_2, \gamma_2), \ldots\}$ known as zeros, where the image's 2-D Z transform (Eqn.(2.69)) is zero, i.e. where $\mathcal{A}(\zeta, \gamma) = 0$. This zero sheet, denoted $\mathrm{Z}\{\mathcal{A}(\zeta, \gamma)\}$, is a single 2-D surface in the 4-D space spanned by the complex coordinates ζ and γ (Lane and Bates, 1987). The discussion in §2.7, which described the expressions that defined the mappings between Z space and Fourier space (Eqns. (2.70) and (2.71)), should indicate that it is a straightforward process to map this zero sheet, obtained from the Z transform, from Z space into Fourier space to obtain the zero sheet that could be obtained from the Fourier transform of the image. It should also be clear from §2.5.3 that in Fourier space the zero sheet of a pixellated image is a periodic repetition of the zero sheet of the equivalent continuous image; provided that the rate of sampling is sufficient to minimise aliasing.

A number of important properties of the surface which is the zero sheet of an image can be derived from the knowledge that the image is compact. The first of these properties is that the zero sheet must extend out to infinity in both the ζ and γ directions since the compact nature of the image implies that the image's Fourier transform and hence Z transform are of infinite extent (Lane and Bates, 1987; Brigham, 1988). The second and third more important properties arise from the fact that the compact nature of the image results in the Z transform (and Fourier transform) of the image being an EFET (see §2.8). A consequence of $\mathcal{A}(\zeta, \gamma)$ being an EFET is that it is completely characterised, to within an arbitrary complex constant, by the set of zeros which form its zero sheet. In addition, the surface which is the zero sheet of the image must be analytic since, as explained by Lane *et al.* (1987), any sudden jump in $\mathrm{Z}\{\mathcal{A}(\zeta, \gamma)\}$ would result in $\mathcal{A}(\zeta, \gamma)$ not being analytic and this would violate the properties that $\mathcal{A}(\zeta, \gamma)$ must possess to be an EFET. In theory it is possible for $\mathrm{Z}\{\mathcal{A}(\zeta, \gamma)\}$ to possess singularities and for $\mathcal{A}(\zeta, \gamma)$ to remain continuous, however the number of conditions that $\mathcal{A}(\zeta, \gamma)$ must satisfy simultaneously for this to occur are almost never met in practice (Bates *et al.*, 1990).

The behaviour of the zero sheet as either ζ or γ become very small or very large can be determined by evaluating the form of the expression for the 2-D Z transform of the image (Eqn.(2.69)) as ζ or γ tend to zero or as ζ or γ tend to infinity respectively. The limits for ζ alone are considered in detail here since the form of the zero sheet at zero or infinity is the same for both coordinates. To determine the form of the zero sheet of the image $a(x, y)$, the 2-D Z transform of the image is rewritten in the form

$$\mathcal{A}(\zeta, \gamma) = \sum_{n=0}^{N-1} \left(a[n, 0] + a[n, 1]\gamma + \cdots + a[n, M-1]\gamma^{M-1} \right) \zeta^n.$$

As $|\zeta|$ tends to infinity it can be seen that all the terms in the above summation become insignificant in comparison to the term containing ζ^{N-1}. In this instance the 2-D polynomial reduces to a 1-D polynomial which has point zeros that are characterised by the pixel values along one edge of the image, i.e. the pixels $a[N-1, 0], a[N-1, 1], \cdots, a[N-1, M-1]$.

Similarly, as $|\zeta|$ tends to zero the only significant term in the summation is that which is scaled by ζ^0, consequently, as $|\zeta| \rightarrow 0$ the zero sheet comprises the point zeros that are characterised by the pixels down the opposite edge of the image, i.e. the pixels $a[0,0], a[0,1], \cdots, a[0,M-1]$. Hence, as $|\zeta|$ or $|\gamma|$ tend to zero or infinity the zero sheet contracts to a set of point zeros and each of these four sets of point zeros are directly related to the pixel values along a corresponding edge of the image (Scivier and Fiddy, 1985; Lane *et al.*, 1987).

From the above discussion it can be seen that in Z space the zero sheet of a 2-D image is an analytic surface in 4-D space. The surface extends to infinity in all directions and as it does so the surface contracts to a set of points characterised by the pixel values along the edges of the image. In Fig. 4.1 a sketch of a typical zero sheet for a 4×4 image is shown to illustrate these properties.

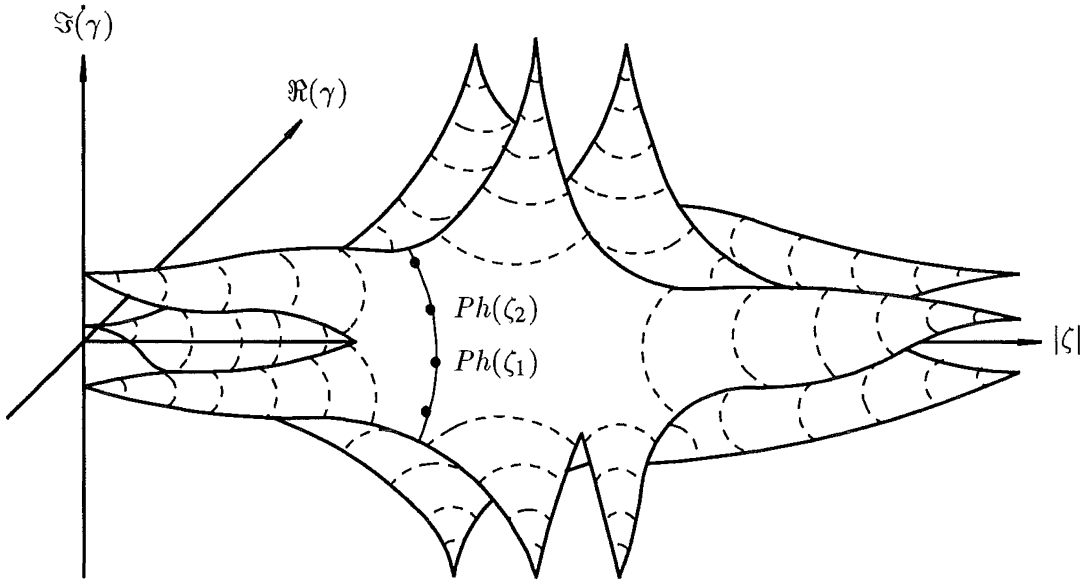

Figure 4.1 A sketch of a typical zero sheet for a 4×4 image showing the shrinking of the surface to a set of point zeros as it tends to zero and infinity. In this figure the fourth coordinate $Ph(\zeta)$ is a parameter around the surface of the sheet. See §4.2 for a complete description of how zero sheet representations of the above form can be produced for an image.

To complete this introduction to zero sheets and their properties the form of the zero sheets of general K-D images, 1-D images, and the relationship between the zero sheets of discrete and continuous 1-D images are now discussed. In the general case where an image exists in K-D space its zero sheet is a $(2K\text{-}2)$-D surface in $2K$-D space (Lane and Bates, 1987). The 1-D imaging situation provides the most straightforward zero sheet since in this context $(2K\text{-}2)=0$. In Z space the zero sheet of a 1-D compact pixellated image therefore consists of a set of point zeros. The compact nature of this 1-D image implies, as with the 2-D case, that the Fourier and Z transform of the image are EFET and as a result the image is completely characterised, to within an arbitrary complex constant, by the finite set of point zeros.

In the situation where the 1-D image is compact and continuous the Hadamard factorisation theorem states, that as a consequence of the image's spectrum being an EFET, the image's spectrum can be expressed in the form of the infinite product

$$A(u) = C \prod_{n=1}^{\infty} \left(1 - \frac{u}{u_n}\right), \tag{4.1}$$

where $A(u)$ is the Fourier transform of the 1-D compact continuous image $a(x)$, C is a complex constant and u_n is the nth complex zero of $A(u)$ (Requicha, 1980; Hurt, 1989). This infinite product, known as the Hadamard product, only exists for 1-D images. The Hadamard product expresses, for 1-D continuous images, the important property of EFET which states that the spectrum of an image is completely characterised, to within a arbitrary complex constant, by its set of zeros. In the situation where the image is in a noisy environment Bates and McDonnell (1986) have shown that, after representing the 1-D compact continuous image $a(x)$ by a finite length Fourier series, it is possible to express the spectrum of the image in the form

$$A(u) = C \prod_{m=2M+1}^{\infty} \left(u - \frac{m}{L}\right)\left(u + \frac{m}{L}\right) \prod_{m=1}^{2M} (u - u_m), \tag{4.2}$$

where C is a complex constant, M is the number of Fourier coefficients that have a magnitude greater than the noise level, L is the extent of the image in image space, the u_m are known as the set of complex zeros of $A(u)$ and the zeros at $\pm\frac{2M+1}{L}, \pm\frac{2M+2}{L}, \ldots$, are known as the set of real zeros since they lie on the real axis. It can be seen that in the noiseless case (i.e. as $M \rightarrow \infty$ in Eqn.(4.2)) Eqn.(4.2) reduces to the Hadamard product (Eqn.(4.1)). It is also interesting to note that if the continuous image is sampled at intervals of $L/2M$ then the zeros of the sampled image's spectrum correspond directly to the $2M$ complex zeros of the image's Fourier transform.

4.1.2 The Uniqueness of the Zero Sheet

No two images have the same zero sheet unless the two images are the same image. This unique relationship between an image and its zero sheet can be obtained from the following simple argument. The Fourier transform and Z transform both define unique invertible mappings between the spatial domain and the frequency domain (see §2.5 and §2.7). Consequently, the Fourier or Z transform of an image is uniquely related to the image from which the spectrum was derived. This spectrum, in the case of a compact image, is an EFET and as a result the Fourier or Z transform of an image is completely characterised, to within the arbitrary complex constant, by the zeros of its transform. These zeros form the zero sheet of the image, therefore, the zero sheet of an image is unique since the set of zeros comprising the zero sheet define a unique spectrum which in turn defines a unique image. It is necessary to note that two images which differ by a complex constant are considered to be the same image since they have the same structure. As described in §3.5.1.1 the image $Ca(\mathbf{x})$, where C is a complex constant, is an image form of $a(\mathbf{x})$.

To uniquely reconstruct a compact pixellated 1-D image from its zeros it is clear from the fundamental theorem of algebra and Hadamard's theorem (Eqn.(4.1)) that all the complex zeros which characterise the image's spectrum are required. In the 2-D case, however, only a subset of the infinity of complex zeros comprising the image's zero sheet are required to obtain a reconstruction (see §4.5). In 2-D a question therefore arises as to how many zeros must be sampled from the zero sheet if the resulting subset of zeros is to uniquely characterise the image from which the zero sheet was derived. The zeros sampled from an image's zero sheet for the purposes of obtaining a reconstruction of the image are known hereafter as *zero samples*. Before introducing the theorems that allow a bound to be set on the number of zero samples that are required to guarantee a unique reconstruction, it is first necessary to define what is meant by the degree of a multivariate polynomial. In 1-D the degree of a polynomial $p(\zeta)$, in which ζ is raised to power $N-1$, is $N-1$. The degree of a 2-D polynomial is defined to be the summation of the degrees in each of the variables (Curtis and Oppenheim, 1987),

hence, the polynomial which is the Z transform of a $N \times M$ pixellated image (Eqn.(2.69)) has degree $(N-1)+(M-1)$.

The number of zero sheet samples that are required to uniquely characterise a 2-D pixellated image can be established by determining the maximum number of zeros two 2-D polynomials can have in common if they are not the same polynomial. The number of common zeros is given by Bezout's theorem (Zakhor and Izraelevitz, 1986; Curtis and Oppenheim, 1987; Hurt, 1989). Bezout's theorem states that two distinct bivariate polynomials of the form

$$P(\zeta,\gamma) = \sum_{r=0}^{R}\sum_{m=0}^{R-r} p[r,m]\zeta^r\gamma^m, \tag{4.3}$$

and

$$Q(\zeta,\gamma) = \sum_{s=0}^{S}\sum_{m=0}^{S-s} p[s,m]\zeta^s\gamma^m, \tag{4.4}$$

have at most RS distinct pairs (ζ,γ) where $P(\zeta,\gamma)=0$ and $Q(\zeta,\gamma)=0$ if neither polynomial can be factorised. Thus, two distinct polynomials of degree R and S which cannot be factorised have at most RS common zeros (Zakhor and Izraelevitz, 1986; Curtis and Oppenheim, 1987; Hurt, 1989). Bezout's theorem therefore indicates that at least $2(N-1)\times 2(N-1)+1 = 4(N-1)^2+1$ zeros are required to uniquely reconstruct an irreducible $N \times N$ image from its zero sheet since the Z transform of the image is a $2(N-1)$ degree polynomial (Zakhor and Izraelevitz, 1986; Curtis and Oppenheim, 1987).

The pixellated images dealt with in most image processing applications have rectangular support, consequently their Z transform polynomials also have rectangular support and not the triangular support implied for the polynomials in Bezout's theorem (see Eqns. (4.3) and (4.4)). It has been shown by that in the situation where the 2-D polynomials have rectangular support the bound in the number of zeros that the two polynomials can have in common is lower than that specified by Bezout's theorem (Zakhor and Izraelevitz, 1986). Zakhor and Izraelevitz (1986) first express the two rectangular polynomials $P(\zeta,\gamma)$ and $Q(\zeta,\gamma)$ in the form

$$\begin{aligned} P(\zeta,\gamma) &= \sum_{r=0}^{R_\zeta}\sum_{m=0}^{R_\gamma} p[r,m]\zeta^r\gamma^m, \\ &= \sum_{r,m=0}^{R_\zeta} P_r(\gamma)\zeta^r, \end{aligned}$$

and

$$\begin{aligned} Q(\zeta,\gamma) &= \sum_{s=0}^{S_\zeta}\sum_{m=0}^{S_\gamma} p[s,m]\zeta^s\gamma^m, \\ &= \sum_{s,m=0}^{S_\zeta} Q_s(\gamma)\zeta^s, \end{aligned}$$

respectively. In the above equations the 1-D polynomials $P_m(\gamma)$ and $Q_m(\gamma)$ are of degree R_γ and S_γ respectively. The number of common zeros can then be determined from an algebraic theorem which states that the two polynomials $P(\zeta,\gamma)$ and $Q(\zeta,\gamma)$ have a common root if

the resultant of the two polynomials $R_{PQ}(\gamma)$, which is defined to be the determinant of the $(S_\zeta + R_\zeta) \times (S_\zeta + R_\zeta)$ matrix,

$$\begin{bmatrix} P_0(\gamma) & P_1(\gamma) & P_2(\gamma) & \cdots & P_{S_\zeta}(\gamma) & 0 & \cdots & 0 \\ 0 & P_0(\gamma) & P_1(\gamma) & P_2(\gamma) & \cdots & P_{S_\zeta}(\gamma) & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & P_0(\gamma) & P_1(\gamma) & P_2(\gamma) & \cdots & P_{S_\zeta}(\gamma) \\ Q_0(\gamma) & Q_1(\gamma) & Q_2(\gamma) & \cdots & Q_{R_\zeta}(\gamma) & 0 & \cdots & 0 \\ 0 & Q_0(\gamma) & Q_1(\gamma) & Q_2(\gamma) & \cdots & Q_{R_\zeta}(\gamma) & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & Q_0(\gamma) & Q_1(\gamma) & Q_2(\gamma) & \cdots & Q_{R_\zeta}(\gamma) \end{bmatrix} \tag{4.5}$$

is zero (Walker, 1950; Zakhor and Izraelevitz, 1986). As $R_{PQ}(\gamma)$ is a polynomial of degree $S_\zeta R_\gamma + R_\zeta S_\gamma$ (the determinant and the polynomial produced when all the elements down the leading diagonal are multiplied together have the same degree (Anton, 1984)) the fundamental theorem of algebra indicates that $R_{PQ}(\gamma)$ must have $S_\zeta R_\gamma + R_\zeta S_\gamma$ zeros and $P(\zeta, \gamma)$ and $Q(\zeta, \gamma)$ must therefore have $S_\zeta R_\gamma + R_\zeta S_\gamma$ common zeros. The above result indicates that $2(N-1)^2+1$ zero samples are required from the zero sheet of a pixellated $N \times N$ image to ensure a unique reconstruction. Curtis and Oppenheim (1987) and Zakhor and Oppenheim (1990) note that if a least squares technique is used to reconstruct the image and more than the minimum $2(N-1)^2+1$ zero samples are used, then the numerical stability of the least squares technique improves with the increase in the number of zero samples.

4.2 DISPLAYING THE ZERO SHEET

The 4-D space in which the zero sheet of a 2-D image exists makes it particularly difficult to visualise the form of the zero sheet or to display the zero sheet using a 2-D medium. In this section a technique which allows a section of a zero sheet to be displayed and which aids the visualisation process is described. The technique based on displaying very small, but nonetheless useful, sections of the zero sheet has been used extensively by the author (Watson and Bones, 1991; Watson *et al.*, 1992; Watson and Bones, 1992), the image processing group at the University of Canterbury (Lane and Bates, 1987; Lane *et al.*, 1987; Bones *et al.*, 1992) and others (Izraelevitz and Lim, 1987). Also alluded to is a second technique that takes advantage of the recent developments in software for the display of multidimensional data and computer hardware to provide a greatly improved means of visualising the form of the zero sheet (Scrase *et al.*, 1992; Satherley, 1994).

4.2.1 Zero Sheet Slices

To obtain a representation of the zero sheet of a compact 2-D image, one of the variables in the 2-D polynomial describing the 2-D image (Eqn.(2.69)), ζ say, is set to an initial constant value. Fixing ζ has the effect of reducing the 2-D polynomial to a 1-D polynomial in γ. The finite set of zeros, which the fundamental theorem of algebra states this resulting 1-D must possess, lie in the γ plane and are a subset of the infinity of zeros that form the image's zero sheet (Lane and Bates, 1987; Lane *et al.*, 1987). A consequence of the analytic nature of the zero sheet of a compact image is that if ζ is now made to follow a continuous path in the ζ plane, the point

zeros of the resulting 1-D polynomials must also trace out continuous loci (hereafter known as *zero tracks*) in the γ plane. If ζ traces out a path that maps out the whole of the ζ plane, a representation of the entire zero sheet of the image, in terms of the ζ path followed, is obtained. In practice it is of course impossible to trace out the entire ζ plane and for this reason a finite length ζ path, which provides a useful amount of information on the zero sheet, is commonly used.

Before describing the form of the paths used by the author it is useful at this point to introduce a limited number of complex analysis techniques which enable the form of the zero sheet and the process by which the zeros are tracked to be understood in a more complete manner. In the situation where sections of the zero sheet of an image are tracked out in the manner described in the previous paragraph it is possible to describe the zero sheet as a complex function of the form $\Omega(\zeta) = \gamma$ (Lane *et al.*, 1987). In general $\Omega(\zeta)$ is a multivalued function (i.e. for a particular value of ζ, $\Omega(\zeta)$ has M zeros where $M > 1$) and to understand the formation of the zero sheet it is useful at this point to think of the ζ plane as a Riemann surface (Lane *et al.*, 1987). A Riemann surface is a generalisation of a complex plane into a surface consisting of a number of sheets which allow a one to one correspondence between the values of the multivalued function and the points on the surface (Titchmarsh, 1939; Fuchs and Levin, 1961; Kreyszig, 1983). Connecting the sheets which form the Riemann surface are singular points or a series of singular points known as branch points or branch cuts respectively. To move from one sheet on the surface to another it is first necessary to pass through one of the singular points. In the case of $\Omega(\zeta)$ these singular points correspond to the values of ζ where two zero tracks come together at a common zero (Izraelevitz and Lim, 1987; Lane *et al.*, 1987). Returning to properties of zero tracks, the analytic nature of the zero sheet implies that if a closed path in the ζ plane, hereafter called a contour, is used to produce a set of zero tracks then these zero tracks must also form closed contours. As a result of the multivalued properties of the function $\Omega(\zeta)$, however, it may be necessary to traverse the ζ contour more than once to ensure that all the γ plane tracks form closed contours (Lane *et al.*, 1987). In more specific terms, if the ζ plane contour crosses a branch cut the path traced out on the Riemann surface moves onto another sheet comprising the surface and it is therefore necessary to continue to move around the ζ plane contour until the path on the Riemann surface returns to the original sheet thereby enabling the zero tracks to form closed contours.

The closed contour used herein for forming zero tracks is given by

$$\zeta = \rho \exp^{i\phi}, \tag{4.6}$$

where ρ is fixed, real, and positive and ϕ is varied by equal angular increments between $-\pi$ and π. In the situation where a $N \times N$ image $f(x, y)$ is under consideration, fixing ζ on the above contour allows the 2-D polynomial describing the image to be rewritten in the form

$$\mathcal{F}(\zeta, \gamma) = A(\zeta)\exp^{i\Psi(\zeta)} \prod_{j=1}^{N-1} (\gamma - \gamma_j(\zeta)), \tag{4.7}$$

where the $\gamma_j(\zeta)$ are the point zeros corresponding to a particular value of ζ and $A(\zeta)$ and $\Psi(\zeta)$ are real. The $\gamma_j(\zeta)$ can be calculated from Eqn.(4.7) by using one of the many standard zero finding algorithms, such as the CPOLY (Jenkins and Traub, 1972) algorithm used by the author (see §5.1). It can also be seen from Eqn.(4.7) that the set of zero tracks for a $N \times N$ image consists of N-1 individual zero tracks. When ζ is moved around the circular contour, it is necessary to ensure that the increments in ϕ are small enough so that each of the zeros in the set $\gamma_j(\zeta_i)$ can be unambiguously linked to their corresponding zeros in the set $\gamma_j(\zeta_{i+1})$. The

zeros in the γ plane of smallest magnitude move least with each increment in ϕ and it has been found that the zero tracks of an image can be most accurately formed by linking together those zeros of smallest magnitude first. An example of a set of zero tracks, for a 5×5 image, formed by moving ζ around the circular contour for $\rho = 0.7$ is displayed in Fig. 4.2. In this figure the beginning of each zero track (corresponding to $\phi = -\pi$) is marked with a bold dot and ϕ is a parameter around the tracks. It can be seen that the set of zero tracks for this 5×5 image consist of four individual zero tracks as predicted by Eqn.(4.7). It can also be seen that the ζ contour must have crossed a branch cut since one of the closed contours in Fig. 4.2 has been formed from more than one zero track. It is also important to note that if two tracks appear to intersect in plots of the type displayed in Fig. 4.2, this does not necessarily mean that the two tracks intersect in 4-D space since intersection only occurs if the ϕ values are the same at the points of intersection on the plot.

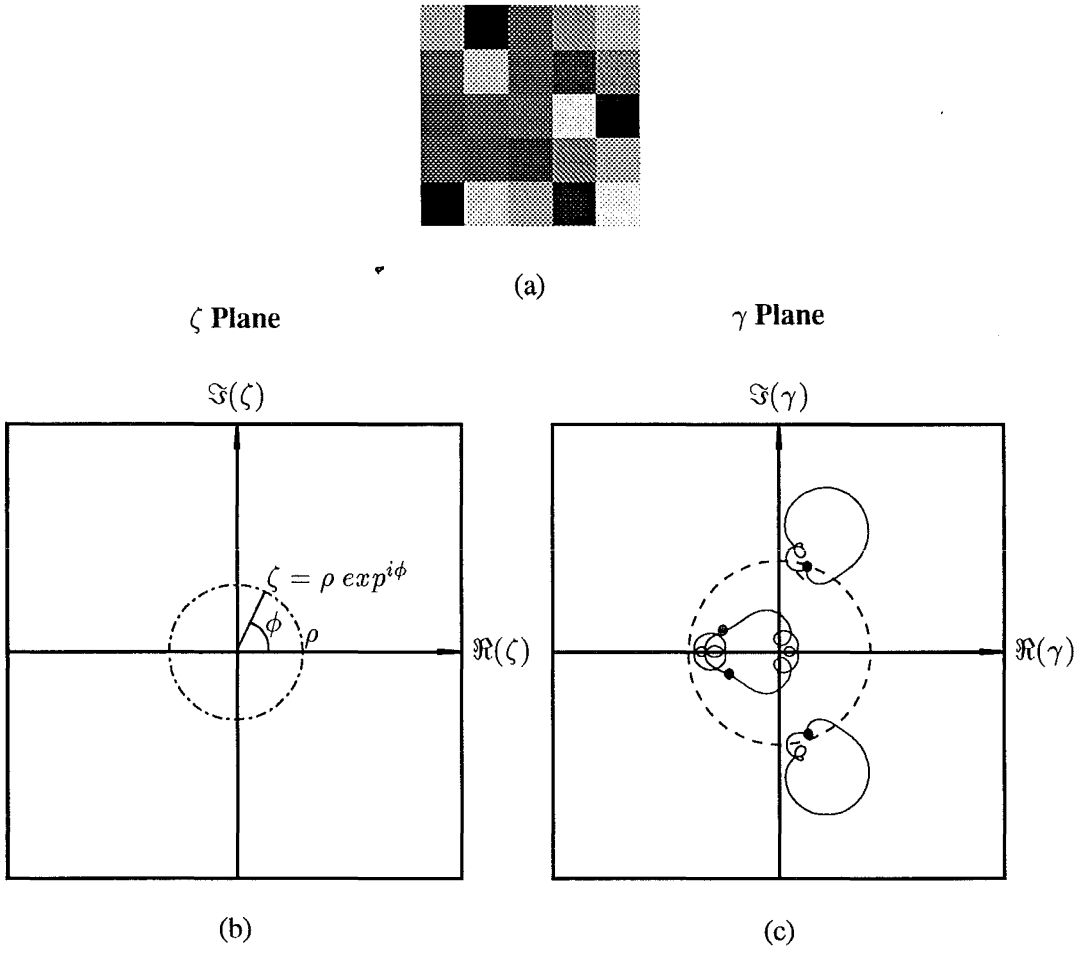

Figure 4.2 (a) a real random 5×5 image, (b) the circular contour $\zeta = \rho\exp^{i\phi}$, and (c) a set of zero tracks of the image in a) for the circular contour with $\rho = 0.7$. The unit circle in the γ plane is shown as a dashed line.

The major advantage of using the circular contour defined by Eqn.(4.6) to produce the sets of zero tracks of an image is that, for each value of ρ, the set of zero tracks so formed represents a slice through a projection of the zero sheet into 3-D space. For this reason, the set of zero tracks produced for a particular ρ value is known as a *slice*. The slices produced for ρ values

ranging between 0 and ∞ can be stacked together to form a representation of the projection of the image's zero sheet into the 3-D space defined by the $\Re(\gamma)$, $\Im(\gamma)$, and ρ coordinates. A series of five slices for the 5×5 image in Fig. 4.2 for $\rho = 0.6, 0.8, 1.0, 1.2, 1.4$ are shown in Fig. 4.3. It is reasonably easy to see how the slices in Fig. 4.3, in addition to a number of other slices for different ρ values, could be stacked together to form a representation of an image's zero sheet such as that originally depicted in Fig. 4.1. This series of slices also effectively demonstrates two of the properties of zero sheets that were first introduced in §4.1.1. These properties are the zero sheets smooth continuous form and the shrinking of the zero sheet towards a set of point zeros as ρ tends to 0 or ∞. Other important properties of zero sheets which can be seen in the slices displayed in Fig. 4.3 are the merging of the closed contours as the ζ contour crosses branch cuts and the real γ axis symmetry that real images must possess if the circular contour defined by Eqn.(4.6) is used to produce the slices (see §4.3.3).

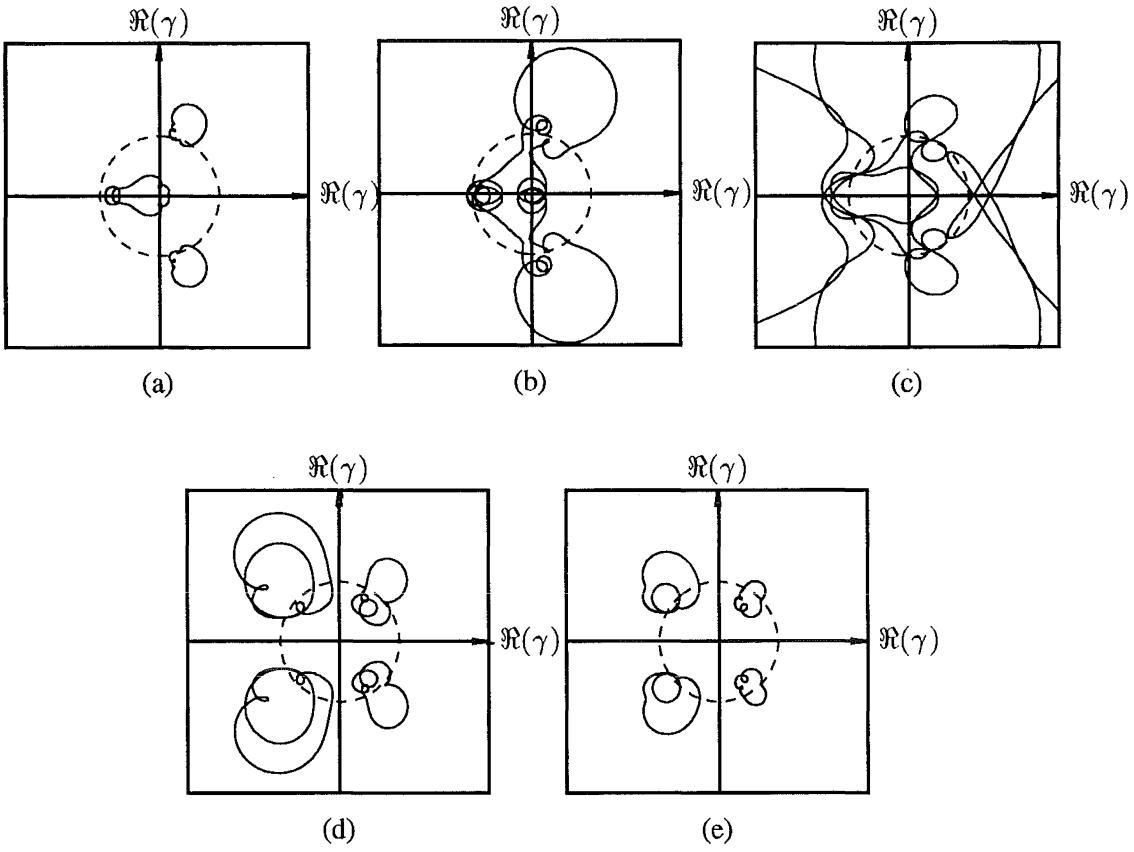

Figure 4.3 A set of slices for the 5×5 image in Fig. 4.2 (a) for the circular contours with the stated ρ values. (a) $\rho = 0.6$, (b) $\rho = 0.8$, (c) $\rho = 1.0$, (d) $\rho = 1.2$, (e) $\rho = 1.4$, and (f) $\rho = 2.0$.

A new method of displaying an image's zero sheet, which provides a means of visualising the form of the zero sheet with greater clarity has been developed recently (Scrase *et al.*, 1992; Satherley, 1994). This new display method produces a representation of the 3-D projection of the zero sheet into (γ,ρ) space, similar to that in Fig. 4.1, by surface rendering a series of the image's zero sheet slices. A series of slices for varying ρ values are first produced by the means outlined above. These slices are then combined to produce a 4-D data set which is passed to the surface rendering program apE (from the Ohio Super Computer Graphics Project at Ohio State University) which tessellates the data and adds to the resulting surface the necessary

shading and reflections required to create a pseudo 3-D representation of the zero sheet. The zero sheet's ϕ dependence is displayed on the 3-D representation by using a cyclic colour map that smoothly changes the colour of the surface with each increment in ϕ (Scrase *et al.*, 1992; Satherley, 1994).

4.3 THE ZERO SHEETS OF USEFUL REDUCIBLE IMAGES

4.3.1 The Uncontaminated Blurred Image's Zero Sheet

In the situation where an image is blurred by a LTI system, the zero sheet of the resulting image has a particularly useful property which can be advantageously used to remove the blurring from an image. Before describing the form of the zero sheet it is first necessary, however, to introduce what turns out to be an exceptionally useful property of multidimensional polynomials and which forms the foundation for all the new zero based deblurring and phase retrieval algorithms discussed in §5.1. In the situation where a polynomial is multivariate the fundamental theorem of algebra and Hadamard factorisation theorem no longer apply, and because of this almost all multivariable polynomials cannot be factorised, i.e. the set of factorable multivariable polynomials is a set of measure zero (Hayes and McClellan, 1982; Curtis and Oppenheim, 1987; Lane *et al.*, 1987). Polynomials which cannot be factorised are known as irreducible polynomials and it can now be seen why pixellated images that cannot be expressed in the form of a convolution (i.e. cannot be expressed as a product of a number of Z transforms) are known as irreducible images. The polynomials which describe the spectrum of a convolved image are one of the few examples of reducible or factorable polynomials. This irreducible nature of compact multidimensional images is also expressed by Osgood's theorem, a multidimensional equivalent of Hadamard's theorem, which states that the spectrum of a multidimensional compact image $a(\mathbf{x})$ can be expressed in the form

$$\mathcal{A}(\zeta) = \prod_{j=1}^{N} [\mathcal{A}_j \exp^{\alpha_j}]^{l_j}, \qquad N \leq \infty$$

where the $\mathcal{A}_j$ are a finite or infinite set of globally irreducible entire functions (not necessarily of exponential type), the $\exp^{\alpha_j}$ terms are convergence factors and the l_j are integers (Scivier and Fiddy, 1985; Hurt, 1989). In the situation where $a(\mathbf{x})$ is an irreducible image there is only one irreducible function $\mathcal{A}_1$ in the Osgood product (i.e. $N = 1$) demonstrating that the polynomial describing the spectrum of the image cannot be factorised (Scivier and Fiddy, 1985)

Returning now to blurred images, it was shown in §2.4 that when a 2-D compact pixellated image is blurred by a LTI system the spectrum of the blurred image, $\mathcal{B}(\zeta, \gamma)$, can be expressed as the product of the spectra, $\mathcal{F}(\zeta, \gamma)$ and $\mathcal{H}(\zeta, \gamma)$, which describe the original image and the psf respectively. In the situation where both the original image and the psf are irreducible it can be seen from the discussion in the previous paragraph that $\mathcal{B}(\zeta, \gamma)$ can only be factorised into the two polynomials which describe $\mathcal{F}(\zeta, \gamma)$ and $\mathcal{H}(\zeta, \gamma)$. As a consequence, $\mathcal{B}(\zeta, \gamma)$ is only zero when $\mathcal{F}(\zeta, \gamma)$ and/or $\mathcal{H}(\zeta, \gamma)$ is zero and the zero set pertaining to $\mathcal{B}(\zeta, \gamma)$ is formed from the union of the zero sets belonging to both $\mathcal{F}(\zeta, \gamma)$ and $\mathcal{H}(\zeta, \gamma)$. These zero sets belonging to $\mathcal{F}(\zeta, \gamma)$ and $\mathcal{H}(\zeta, \gamma)$ describe two separate zero sheets, $\mathrm{Z}\{\mathcal{F}(\zeta, \gamma)\}$ and $\mathrm{Z}\{\mathcal{H}(\zeta, \gamma)\}$, which must therefore coexist in Z space forming the zero sheet of the blurred image $\mathrm{Z}\{\mathcal{B}(\zeta, \gamma)\}$. In more explicit terms,

$$\mathrm{Z}\{\mathcal{B}(\zeta, \gamma)\} = \mathrm{Z}\{\mathcal{F}(\zeta, \gamma)\} \cup \mathrm{Z}\{\mathcal{H}(\zeta, \gamma)\}, \tag{4.8}$$

where $\cup$ denotes the set union operation. This property of $Z\{\mathcal{B}(\zeta,\gamma)\}$ is demonstrated pictorially in Fig. 4.4 where it can be seen that the blurred image's zero sheet slice for $\rho = 0.8$ is indeed composed of the two separate zero sheets slices belonging to the original image and the psf.

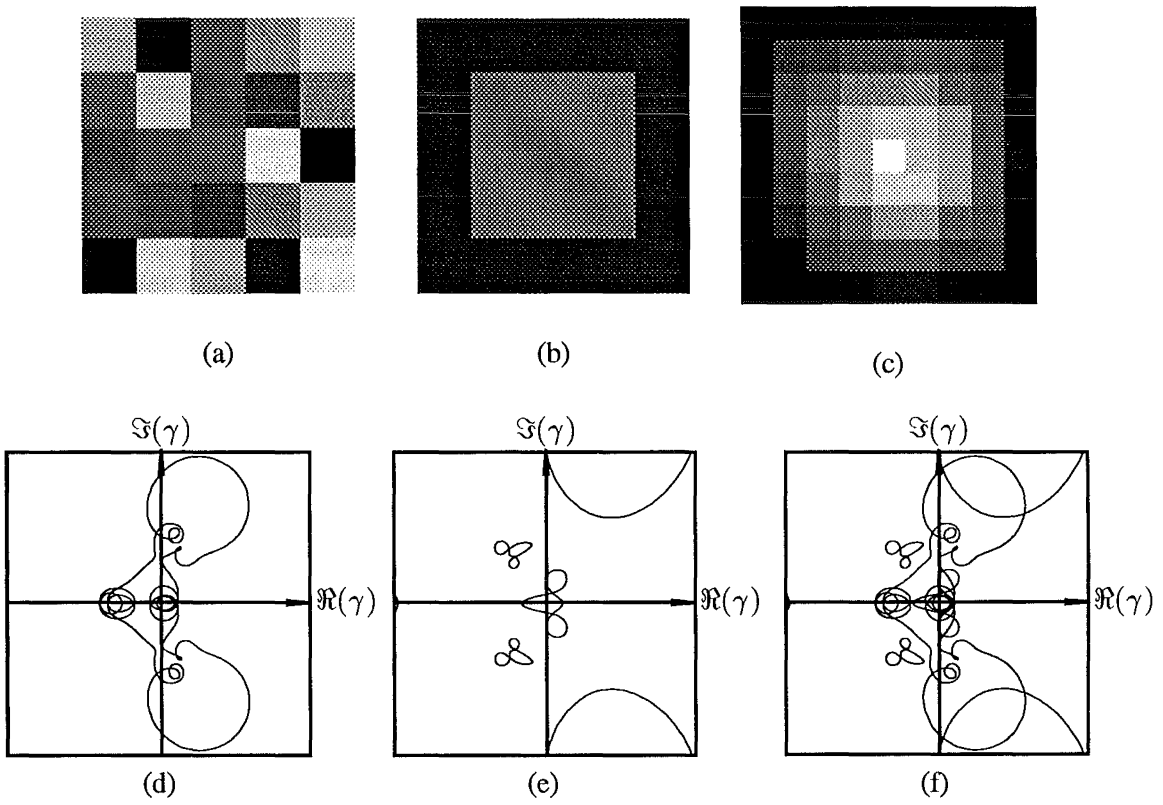

Figure 4.4 The zero tracks of a convolution. (a) the original 5×5 image, (b) the psf, (c) the convolution of the image in (a) and the psf in (b), (d) the $\rho = 0.8$ zero sheet slice belonging to the image, (e) the $\rho = 0.8$ zero sheet slice belonging to the psf, and (f) the $\rho = 0.8$ zero sheet slice belonging to the convolution in (c).

It is the above property of the zero sheets of blurred 2-D images, and in general those of multidimensional convolutions, which is the basis for the new deconvolution algorithm described in §5.1 and the zero based algorithms recently proposed or developed by others (Lane and Bates, 1987; Bates *et al.*, 1990; Parker, 1992; Satherley and Bones, 1993; Ghiglia *et al.*, 1993). As the zero sheet of a blurred image is a 2-D surface in the 4-D space comprising the two separate 2-D surfaces belonging to each component in the convolution, it would appear that it should be possible to separate the two surfaces and reconstruct the image from samples from one surface and the psf from samples of the other. Deblurring an image by this method presupposes of course that is is possible to recognise and separate the individual surfaces comprising $Z\{\mathcal{B}(\zeta,\gamma)\}$. The properties of zero sheets and multidimensional surfaces in multidimensional spaces, which are now outlined, indicate that this separation is indeed feasible.

The first property is derived from a theorem concerned with high dimensional geometry which describes the form of the intersections of multidimensional surfaces in multidimensional spaces. The theorem states that two different manifolds (a manifold is any smooth nonsingular K-D surface situated in Euclidean space of some, perhaps much higher, dimension)

of dimension M and P which exist in a N-D space intersect on manifolds of dimension $M + P - N$ within this N-D space, e.g. two different lines in 2-D space intersect at points since $1 + 1 - 2 = 0$ (Dubrovin *et al.*, 1985). The two separate 2-D surfaces in 4-D space representing the band-limited image must therefore intersect at points in this 4-D space (i.e. $2 + 2 - 4 = 0$) (Eves, 1965; Lane *et al.*, 1987; Lane, 1988). The number of points where these two surfaces intersect can be determined from the extension of Bezout's theorem (see §4.1.2) presented by Zakhor and Izraelevitz (1986), which also implies that the two surfaces must intersect at points since the number of points two different zero sheets can have in common is necessarily finite. The second property which aids the separation, is that the two surfaces must be continuous since they are the zero sheets of compact images and consequently must be analytic. It is therefore possible to construct a zero sheet separation algorithm, as Lane *et al.* (1987) and more recently Ghiglia *et al.* (1993) have done, which follows one of the two surfaces and applies the continuity condition at the points of intersection to determine to which surface the subsequent zeros belong. It is worth noting that it is of course impossible to determine which one of the two surfaces represents the original image and which represents the psf, however in many cases the physical characteristics of the imaging situation should indicate which component is which.

The above properties of the zero sheet of compact images allow any noiseless blurred image of dimension two or greater to be automatically deblurred. In the 1-D case, however, it is not possible to separate the zero sheet belonging to $\mathcal{B}(\zeta)$ into the two distinct zero sheets belonging to $\mathcal{F}(\zeta)$ and $\mathcal{H}(\zeta)$ without having prior knowledge of either $\mathcal{F}(\zeta)$ or $\mathcal{H}(\zeta)$. The reason why it is not possible to uniquely separate the zero sheet of $\mathcal{B}(\zeta)$ into those of $\mathcal{F}(\zeta)$ and $\mathcal{H}(\zeta)$ is as follows. In 1-D the fundamental theorem of algebra implies that $\mathcal{B}(\zeta)$, $\mathcal{F}(\zeta)$, and $\mathcal{H}(\zeta)$ are all reducible polynomials, therefore, $\mathcal{B}(\zeta)$ can be arbitrarily expressed in the form of any one of its many possible factorisations. Each of the factors of $\mathcal{B}(\zeta)$ has a set of zeros which are a subset of the zeros of $\mathcal{B}(\zeta)$ and there can be no link between these individual zeros of $\mathcal{B}(\zeta)$ (i.e. they cannot form smooth surfaces representing irreducible polynomials as in the 2-D case) since this would preclude most of the factorisations that $\mathcal{B}(\zeta)$ must have to satisfy the fundamental theorem of algebra. For this reason, it is not possible to distinguish which sets of zeros belong to $\mathcal{F}(\zeta)$ and $\mathcal{H}(\zeta)$, which is required to automatically deblur the blurred image, without prior knowledge of one of the zero sets of $\mathcal{F}(\zeta)$ or $\mathcal{H}(\zeta)$. This property of the zero sheets of 1-D images is particularly relevant to the phase retrieval problem which was introduced in §3.5.

4.3.2 The Zero Sheet of an Uncontaminated Power Spectrum

It can be shown, in a manner similar to that used in §4.3.1 for blurred images, that the zero sheet of a power spectrum is composed of the two separate zero sheets that belong to the original image and the image's conjugate reflection respectively. This property is particularly useful since it should be possible to use zero sheet separation to solve the phase retrieval problem. That is, in the case where only the Fourier magnitude of an object of interest can be measured, such as in the antenna aperture imaging (McCormack *et al.*, 1990) and crystallography problems (Millane, 1990), it should be possible to use zero sheet separation to recover phase. In recent years a number of researchers have developed algorithms, which use this very zero sheet property, to solve the phase retrieval problem (Izraelevitz and Lim, 1987; Lane *et al.*, 1987; Parker and Bones, 1992).

In order to demonstrate that the zero sheet of a power spectrum does indeed consist of the two separate zero sheets pertaining to the image and its conjugate reflection it is necessary to

return to some of the Fourier and Z transform properties introduced in §2.5 and §2.7. It was shown in §2.4 that the measured power spectrum (i.e. that on the real (u,v) plane) of an image is equal to the spectrum of the image's auto-correlation (Eqn.(2.58)). In the case where the original image $f(x,y)$ is both compact and irreducible, the polynomial

$$\mathcal{Z}_{ff}(\zeta,\gamma) = \mathcal{F}(\zeta,\gamma)\mathcal{F}^*(\frac{1}{\zeta^*},\frac{1}{\gamma^*})\zeta^{N-1}\gamma^{M-1},$$

which describes the Z spectrum of the auto-correlation (see §2.7), can only be factorised into the two irreducible polynomials representing the image and its conjugate reflection. The fact that $\mathcal{Z}_{ff}(\zeta,\gamma)$ can only be factorised into two irreducible components is a consequence of the properties of 2-D polynomials introduced in §4.3.1. As a result of $\mathcal{Z}_{ff}(\zeta,\gamma)$ only being separable into two factors it can be seen from §4.3.1 that the zero sheet of the power spectrum can be expressed in the form

$$\mathsf{Z}\{\mathcal{Z}_{ff}(\zeta,\gamma)\} = \mathsf{Z}\{\mathcal{F}(\zeta,\gamma) \cup \mathsf{Z}^*\{\mathcal{F}(\frac{1}{\zeta^*},\frac{1}{\gamma^*})\}, \tag{4.9}$$

and consequently the zero sheet of an image's power spectrum comprises the two separate zero sheets belonging to the image and its conjugate reflection (see Fig. 4.6). For the reasons outlined in §4.3.1 it is possible to separate the two surfaces thereby enabling a reconstruction of the image and its conjugate reflection to be obtained. As with the deblurring problem it is impossible to know without prior knowledge which zero sheet belongs to the image and which belongs to the conjugate reflection.

The concept of zero sheets and the irreducible nature of multidimensional polynomials provides the most straightforward method of demonstrating that the phase retrieval problem does not have a unique solution in 1-D but does in 2 or more dimensions. Basically, the phase retrieval problem has a unique solution in spaces with a dimensionality of two or greater since in two or more dimensions $\mathcal{F}_{ff}(\zeta)$ cannot be factorised beyond its irreducible components $\mathcal{F}(\zeta)$ and $\mathcal{F}(\frac{1}{\zeta^*})$, consequently, $\mathsf{Z}\{\mathcal{F}_{ff}(\zeta)\}$ consists of the two separate zero sheets pertaining to $\mathcal{F}(\zeta)$ and $\mathcal{F}(\frac{1}{\zeta^*})$. In the 1-D situation, as described in §4.3.1 for the deconvolution situation, $\mathcal{F}_{ff}(\zeta)$ can be factorised into any of its many possible factorisations and therefore in 1-D the phase retrieval problem does not have a unique solution. The possibility of obtaining a unique solution to the phase retrieval problem in $\geq$ 2-D was originally demonstrated, in terms of the properties of the power spectra zero sheets, by Lane *et al.* (1987) and Izraelevitz and Lim (1987).

4.3.3 Zero Sheet Symmetries

In this section the symmetry that exists in the zero sheets of both reducible and irreducible images is discussed. The properties described are not merely introduced from the point of view of completeness but also because some of the properties are used to advantage in the solution of the phase retrieval problem and could be used in future deblurring algorithms to aid the zero sheet separation.

The first form of zero sheet symmetry to be introduced is the symmetry which all zero sheets belonging to real images must possess. In the case of real images $f(x,y) = f^*(x,y)$, consequently the Z transform properties in Table 2.2 indicate that $\mathcal{F}(\zeta,\gamma) = \mathcal{F}^*(\zeta^*,\gamma^*)$. If $\mathcal{F}(\zeta,\gamma)$ equals zero at (ζ_1,γ_1) then due to the relationship outlined above it must also be zero at (ζ_1^*,γ_1^*). In the situation where the circular contour defined by Eqn.(4.6) is used to produce

a slice of an image's zero sheet, the set of zero tracks are symmetric in the real γ axis if the ϕ dependence of the tracks is ignored (i.e. if $\mathcal{F}(\rho \exp^{i\phi_1}, \gamma_1) = 0$ then, when $\zeta = \rho \exp^{-i\phi_1}$, one of $\mathcal{F}(\zeta, \gamma)$'s zeros is at γ_1^*). This symmetry is clearly evident in many of the previous figures in this section, for example Fig. 4.4, and was first alluded to in §4.2.1. This symmetry is only evident in a set of slices produced from the closed circular contour defined by Eqn.(4.6) since no other contour has the property that if $\rho \exp^{i\phi}$ is on the contour then so is $\rho \exp^{-i\phi_1}$, for all possible ϕ values.

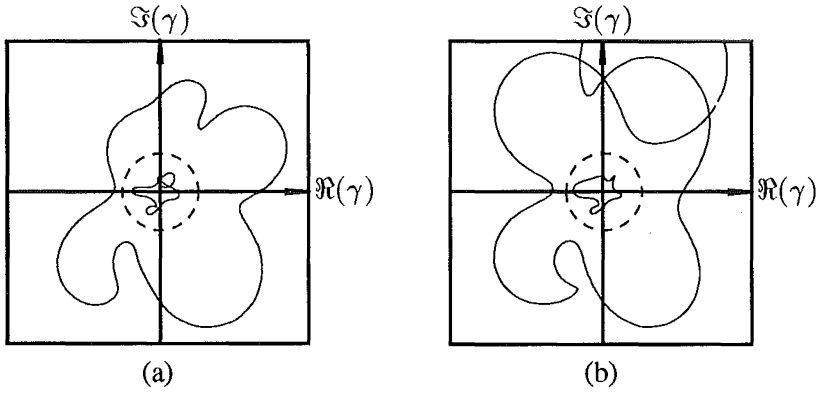

Figure 4.5 The symmetry through the unit circle at ρ, (a) $\rho = 0.75$, and $\frac{1}{\rho}$, (b) $\rho = 1.33$, in the zero sheet slices of an image that has conjugate reflection symmetry. The unit circle is shown as a dashed line.

The second type of zero sheet symmetry to be discussed arises with images that have conjugate reflection symmetry given by $f(x, y) = f^*(-x, -y)$. Using the Z transform properties in Table 2.2 it can be seen that $\mathcal{F}(\zeta, \gamma) = \zeta^{N-1}\gamma^{M-1}\mathcal{F}^*(\frac{1}{\zeta^*}, \frac{1}{\gamma^*})$ and therefore if $\mathcal{F}(\zeta_1, \gamma_1)$ is zero, then so is $\mathcal{F}(\frac{1}{\zeta_1^*}, \frac{1}{\gamma_1^*})$. If sets of zero tracks are produced for the closed circular contours $\zeta = \rho \exp^{i\phi}$ and $\zeta = \frac{1}{\rho} \exp^{i\phi}$ the zero tracks for the ρ slice are reflections through the unit circle of those at $\frac{1}{\rho}$, since if $\mathcal{F}(\rho \exp^{i\phi}, \gamma_1) = 0$ then $\mathcal{F}(\frac{1}{\rho} \exp^{i\phi}, \frac{1}{\gamma_1} \exp^{iPh(\gamma_1)})$ must also be zero as is shown in Fig. 4.5. This form of symmetry is known hereafter as *partial unit circle symmetry*.

Another useful form of reducible image which produces zero sheets with partial unit circle symmetry is that of an auto-correlation. The Z spectrum of the auto-correlation of $f(x, y)$ consists of the two components $\mathcal{F}(\zeta, \gamma)$ and $\mathcal{F}^*(\frac{1}{\zeta^*}, \frac{1}{\gamma^*})$, and since $\mathrm{Z}\{\mathcal{F}(\zeta, \gamma)\}$ has partial unit circle symmetry with that of $\mathrm{Z}\{\mathcal{F}^*(\frac{1}{\zeta^*}, \frac{1}{\gamma^*})\}$ the zero sheet of the auto-correlation must also possess partial unit circle symmetry as can be seen in Fig. 4.6. This form of symmetry is most evident for the slice produced using the unit circle because in this situation $\rho = 1/\rho$ and both sets of tracks which are related via this symmetry are on the same slice. In the 1-D case this symmetry can be used to reduce the number of possible solutions to the phase retrieval problem (i.e. it is known that $\mathcal{F}(\zeta)$ cannot simultaneously possess the factors $(\zeta - \zeta_1)$ and $(\zeta - \frac{1}{\zeta_1^*})$) and in $\geq$ 2-D this property could be used to aid the separation of zero sheets under contaminated conditions (the effects of noise are described in §4.4). It is also worth noting that the two component zero sheets of an image's auto-correlation can only intersect on the unit circle for the slice produced for $\rho = 1$ since for the zero sheets to intersect $\mathcal{F}(\zeta, \gamma) = \mathcal{F}^*(\frac{1}{\zeta^*}, \frac{1}{\gamma^*}) = 0$ and this only occurs when $\zeta = \frac{1}{\zeta^*} = 1 \exp^{i0}$ and $\gamma = \frac{1}{\gamma^*} = 1$ (Lane *et al.*, 1987).

The final type of symmetry to be considered arises when an image has the form $f(x, y) = f(|x|, |y|)$ which is known hereafter as radial symmetry. Images of this form arise regularly in image deblurring problems because the psfs which represents out-of-focus blur and certain

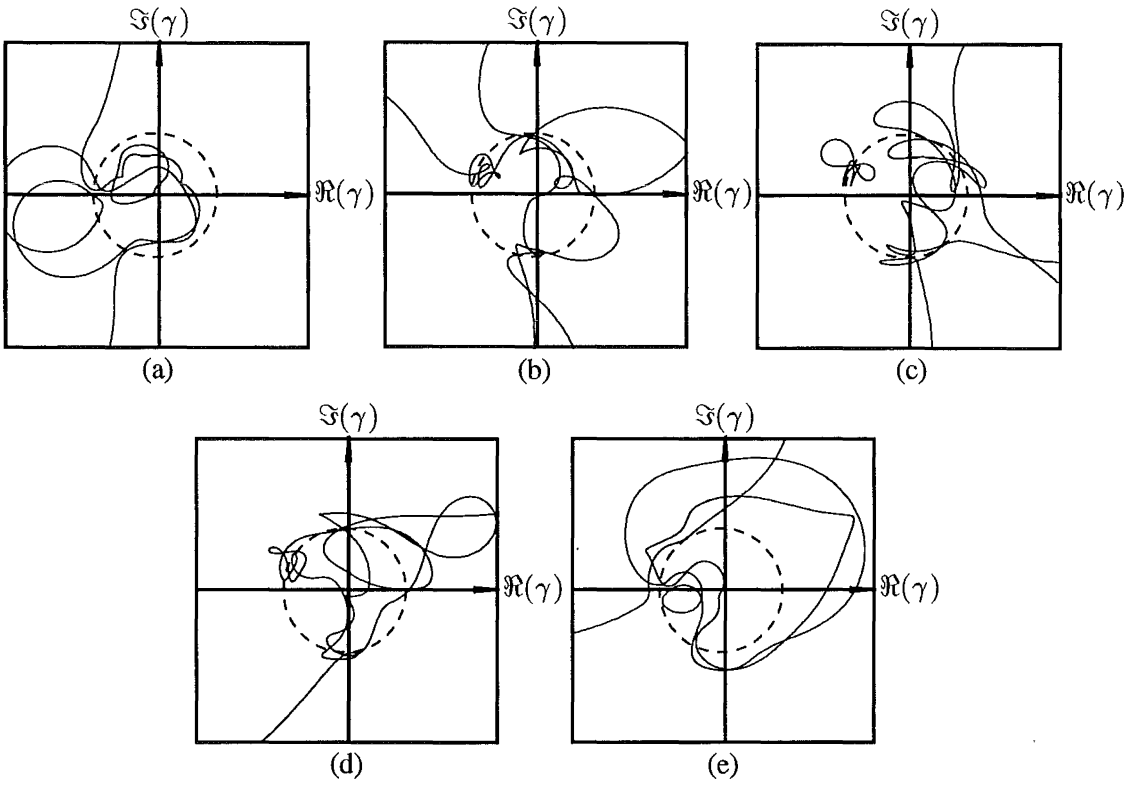

Figure 4.6 Slices from the zero sheet of the power spectrum of a random complex 4×4 image for the ρ values stated. (a) $\rho = 0.5$, (b) $\rho = 0.75$, (c) $\rho = 1.0$, (d) $\rho = 1.33$, and (e) $\rho = 2.0$. The unit circle is shown as a dashed line.

other blurs arising in optical systems have a form which approximates radial symmetry. When an image has radial symmetry the Z transform properties in Table 2.2 indicate that the Z spectrum of the image has the property that

$$\mathcal{F}(\zeta,\gamma) = \mathcal{F}(\frac{1}{\zeta},\frac{1}{\gamma})\zeta^{N-1}\gamma^{M-1} = \mathcal{F}(\zeta,\frac{1}{\gamma})\gamma^{M-1} = \mathcal{F}(\frac{1}{\zeta},\gamma)\zeta^{N-1}.$$

At each value of ρ the slice produced must therefore consist of two sets of tracks, with one set being the reflection of the other set in the unit circle (i.e. the symmetry that arises at $\rho = 1$ in the case of the auto-correlation occurs at all values of ρ for images with radial symmetry). In addition, on ignoring the ϕ dependence the set of zero tracks at ρ have the same form as those at $1/\rho$. This form of symmetry is known hereafter as *full unit circle symmetry*. The zero sheet symmetry that arises with images that are radially symmetric is pictorially demonstrated in Fig. 4.7 where the slices of a convolution (the same image and radially symmetric psf used to produce Fig. 4.4) are shown. The full unit circle symmetry of the psf's zero sheet can be clearly seen in Fig. 4.7 whilst the zero sheet of the image, which does not possess unit circle symmetry, can also be easily distinguished. The figure also demonstrates that if an image has more than one of the properties described above which result in some form of symmetry, the zero sheet of the image must also possess more than one form of symmetry. In the situation where a radially symmetric psf blurs an image it may be possible to use the symmetry property to improve the deblurring algorithm described in §5.1.

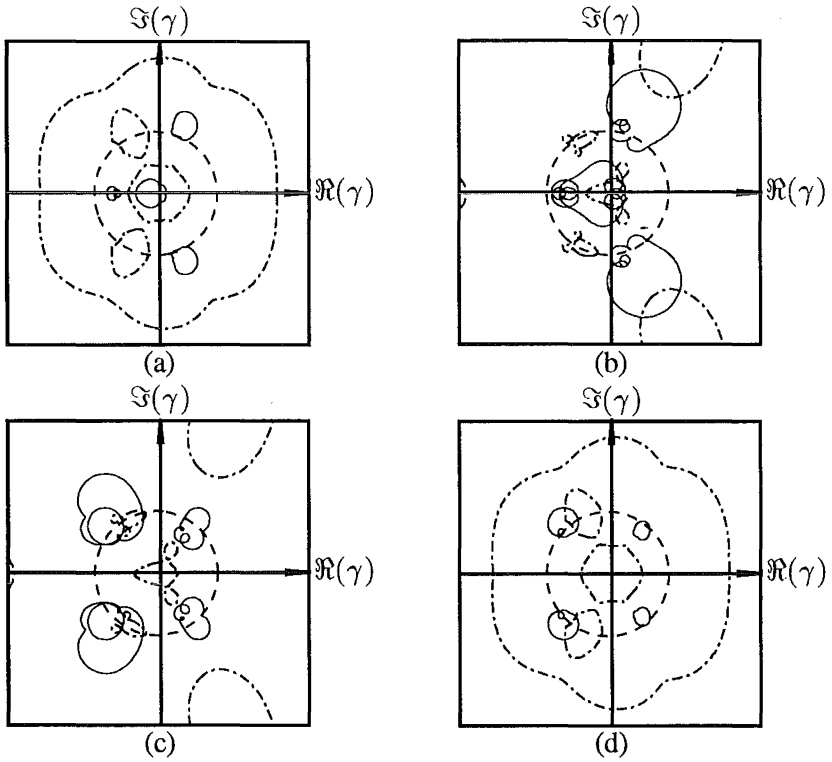

Figure 4.7 Symmetry of the zero sheet of an image that has radial symmetry. Shown is the zero sheet slices for the convolution in Fig. 4.4 where the radially symmetric psf is represented by the dash-dot lines. (a) $\rho = 0.5$, (b) $\rho = 0.75$, (c) $\rho = 1.0$, (d) $\rho = 1.33$, and (e) $\rho = 2.0$. The unit circle is shown as a dashed line.

4.4 THE EFFECT OF CONTAMINATION ON ZERO SHEETS

4.4.1 Wide-Band Contamination

In the situation where a pixellated image is contaminated by an appreciable amount of wide-band additive contamination, the zero sheet of the contaminated image is recognisably similar to that of the uncontaminated image, with the contamination causing smooth perturbations on the zero sheet belonging to what was the uncontaminated image. The smooth nature of the perturbations on the zero sheets of contaminated images is to be expected from the analytic nature of the zero sheet. That is, if an infinitesimal amount of contamination is gradually added to the image, the zero sheet of the image must transform from that of the uncontaminated image in a smooth continuous manner, since the zero sheet of the image must remain analytic at all times. In the situation where the contamination has a bandwidth sufficient to perturb all the image's spectral components within the image's bandwidth, the entire Z spectrum of the image is perturbed and as a consequence the entire zero sheet of the image can be expected to be perturbed. As the level of additive contamination increases, the analytic nature of the image's zero sheet and Z spectrum indicate that there must be a corresponding increase in the size of the perturbations on the image's zero sheet (Lane and Bates, 1987). The effects of wide-band contamination on the zero sheet of a compact pixellated image is shown in Fig. 4.8 for four different levels of noise. Each of the ensembles of slices shown in Fig. 4.8 has been formed by superimposing the $\rho = 0.7$ zero sheet slices of ten differently contaminated (the contamination distributions and SNRs on the ten images in each ensemble are the same) versions of the same 5×5 image. In Fig. 4.2 the $\rho = 0.7$ zero sheet slice pertaining to the uncontaminated version

of the image is displayed. Figure 4.8 confirms that the additive contamination does indeed smoothly perturb the entire zero sheet of the image, with the size of the perturbations increasing as the level of noise increases.

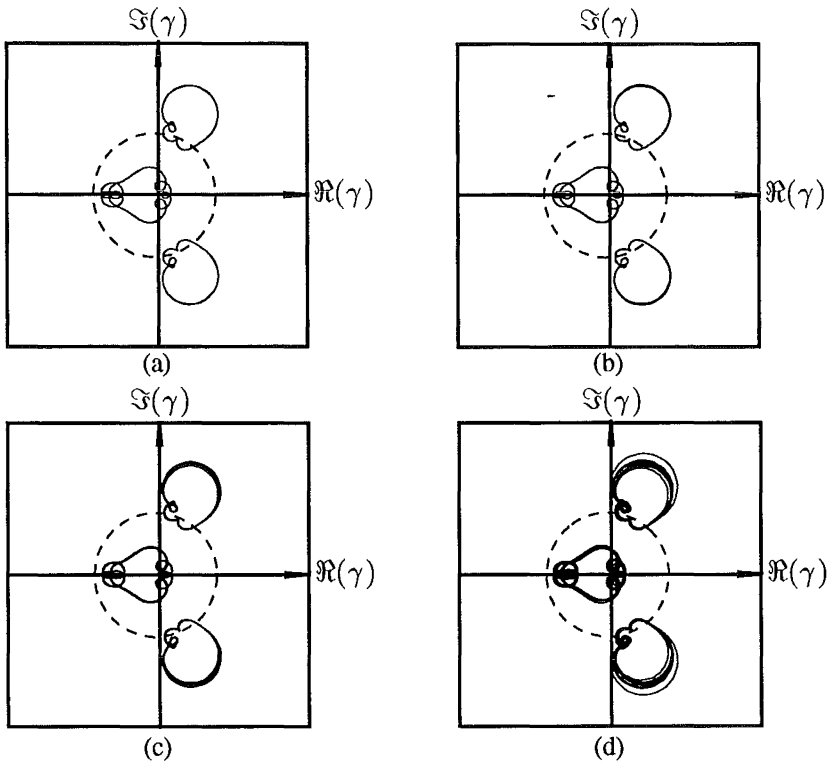

Figure 4.8 The effects of wide-band contamination on an images zero sheet for the slice at $\rho = 0.7$ for an image SNR of (a) 60dB, (b) 50 dB, (c) 40 dB , and (d) 30 dB. The gradual thickening of the lines comprising each ensemble and their eventual separation as the SNR decreases should be noted.

In addition to the properties of the zero sheets of contaminated images described above, the sets of slices in Fig. 4.8 also demonstrate that in Z space the absolute zero sheet perturbation, caused by the additive contamination, tends to increase as $|\gamma|$ increases. It is postulated that the reason for this behaviour is that the mappings defined by Eqns. (2.70) and (2.71) have the effect of compressing half the 4-D *volume* in Fourier space to within the 4-D unit *sphere* in Z space and hence the effects of the contamination are also compressed within this region. At first appearance it looks as if this property could be advantageously used to improve the quality of the reconstructions obtained when the linear equation method (see §4.5.2) is used to reconstruct an estimate of an image from its zero sheet. The reason for this is that, if it is possible to predict which regions of the zero sheet are affected least by the contamination (i.e. possibly those regions with small $|\gamma|$) it would be possible to select zero samples from these regions and thereby obtain reconstructions that are superior to those obtained from zero samples from other regions. It is worth noting that if the Fourier reconstruction method (see §4.5.1) is used to obtain a reconstruction no improvement could be expected because the method has rather tight constraints on where the zero samples can be taken from the zero sheet (i.e. it is unlikely that all the required samples have small $|\gamma|$), and the zero samples must be mapped back onto the real Fourier plane via Eqns. (2.70) and (2.71) which in effect reverses the noise compression.

As reported in Watson *et al.* (1992), the author originally believed that selecting zeros of the smallest $|\gamma|$ would result in the best reconstructions, however on closer examination it appears

that the relative zero sheet perturbation, and not the absolute perturbation, must decrease with decreasing $|\gamma|$ if selecting zero samples of small $|\gamma|$ is to provide superior reconstructions with the linear equation method. The reason why the relative perturbation is most important is because during the process of solving a system of linear equations each equation within the set of simultaneous equations is usually multiplied by a constant at some stage to allow the system to be reduced. This multiplication operation scales the absolute magnitude of the contamination on each of the equations, consequently it is the relative size of the contamination on the coefficients (i.e. the ratio of the coefficient to the contamination on that coefficient) in each equation which dictates the final accuracy of the solution. The coefficients in the system of linear equations used to reconstruct an image from its zero sheet are derived from the zero samples used to construct the system of equations and therefore the relative amount of contamination on these zero samples dictates the reconstruction quality.

To determine whether zero samples with small $|\gamma|$ are perturbed less in relative terms than those with large $|\gamma|$, when the image is corrupted with additive contamination, the following analysis was performed. Consider two unperturbed samples of an image's Fourier space zero sheet, v_1 and v_2, which have the property that they are reflections of each other in the real v axis (i.e. $v_1 = v_2^*$) as shown in Fig. 4.9. If these two zeros are mapped into Z space by using Eqn.(2.71) the resulting two zeros, γ_1 and γ_2, lie in the γ plane and have the properties: $|\gamma_1| > |\gamma_2|$, $|\gamma_1| = |\frac{1}{\gamma_2}|$, and $Ph(\gamma_1) = Ph(\gamma_2)$, as is shown in Fig. 4.9. To determine whether γ_1 is perturbed more by wide-band contamination than γ_2, an equal amount of contamination, Δv, is added to both v_1 and v_2 (in a manner that ensures that the perturbed zeros are still reflections of each other in the real v axis) to give the perturbed Fourier space variables

$$\tilde{v_1} = v_1 + \Delta v \tag{4.10}$$

$$\tilde{v_2} = v_2 + \Delta v^* \tag{4.11}$$

If in Z space γ_1 is perturbed more in relative terms than γ_2 by the Fourier space contamination, then the expression

$$\begin{aligned} |RC_{\gamma_1\gamma_2}| &= \left|\frac{\Delta\gamma_1/\gamma_1}{\Delta\gamma_2/\gamma_2}\right| \\ &= \frac{|\tilde{\gamma_1} - \gamma_1||\gamma_2|}{|\tilde{\gamma_2} - \gamma_2||\gamma_1|}, \end{aligned} \tag{4.12}$$

which is the ratio of the relative perturbation on γ_1 to that on γ_2 (known hereafter as the relative contamination ratio), is greater than one. In the above expression the perturbed Z space variables are denoted by $\tilde{\gamma_1}$ and $\tilde{\gamma_2}$, and $\Delta\gamma_1$ and $\Delta\gamma_2$ are the amounts γ_1 and γ_2 are perturbed by the contamination respectively. Using Eqn.(2.71) the perturbed Z space variables can be expressed in the form

$$\begin{aligned} \tilde{\gamma_1} &= \exp^{-\Omega(\Im(v_1)+\Im(\Delta v))}\exp^{i\Omega(\Re(v_1)+\Re(\Delta v))} \\ &= \exp^{-\Omega\Im(v_1)}\exp^{-\Omega\Im(\Delta v)}\exp^{i\Omega\Re(v_1)}\exp^{i\Omega\Re(\Delta v)} \end{aligned} \tag{4.13}$$

$$\begin{aligned} \tilde{\gamma_2} &= \exp^{-\Omega(\Im(v_2)+\Im(\Delta v^*))}\exp^{i\Omega(\Re(v_2)+\Re(\Delta v^*))} \\ &= \exp^{-\Omega\Im(v_2)}\exp^{-\Omega\Im(\Delta v^*)}\exp^{i\Omega\Re(v_2)}\exp^{i\Omega\Re(\Delta v^*)}, \end{aligned} \tag{4.14}$$

where $\Omega = 2\pi\epsilon$. As $v_2 = {v_1}^*$ it is possible to rewrite Eqn.(4.14) in the form

$$\tilde{\gamma_2} = \exp^{\Omega\Im(v_1)}\exp^{\Omega\Im(\Delta v)}\exp^{i\Omega\Re(v_1)}\exp^{i\Omega\Re(\Delta v)}.$$

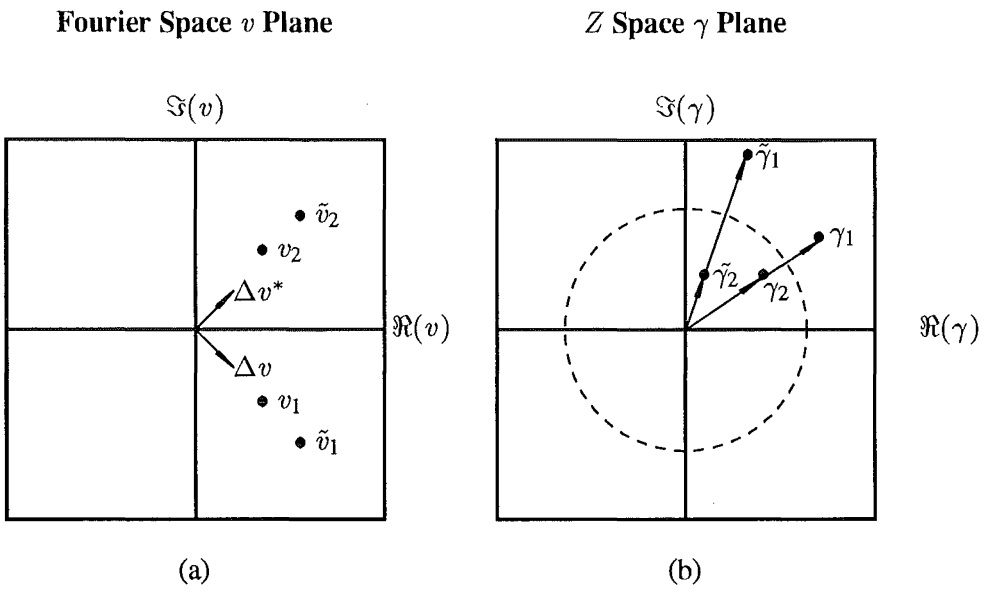

Figure 4.9 (a) Two Fourier space zeros v_1 and v_2 which are conjugates of each other, and (b) the equivalent zeros mapped into Z space. The unit circle is shown as a dashed line.

For reasons of clarity it is now assumed that the inter-pixel spacing $\epsilon = \frac{1}{2\pi}$ and therefore $\Omega = 1$.

Making use of the identities $r\exp^{i\theta} = r\cos(\theta) + ir\sin(\theta)$ and $\cos^2(\theta) + \sin^2(\theta) = 1$, it is possible to express the magnitude of the γ space perturbations, $|\Delta\gamma_1|$ and $|\Delta\gamma_2|$, in the form

$$\begin{aligned} |\tilde{\gamma}_1 - \gamma_1| &= |\exp^{-\Im(v_1)}\exp^{i\Re(v_1)}[\exp^{-\Im(\Delta v)}\exp^{i\Re(\Delta v)} - 1]| \\ &= \exp^{-\Im(v_1)}[\exp^{-2\Im(\Delta v)} - 2\exp^{-\Im(\Delta v)}\cos(\Re(\Delta v)) + 1]^{1/2} \end{aligned} \quad (4.15)$$

and

$$\begin{aligned} |\tilde{\gamma}_2 - \gamma_2| &= |\exp^{\Im(v_1)}\exp^{i\Re(v_1)}[\exp^{\Im(\Delta v)}\exp^{i\Re(\Delta v)} - 1]| \\ &= \exp^{\Im(v_1)}[\exp^{2\Im(\Delta v)} - 2\exp^{\Im(\Delta v)}\cos(\Re(\Delta v)) + 1]^{1/2}. \end{aligned} \quad (4.16)$$

Substituting Eqn.(4.15), Eqn.(4.16) and the mapping expressions which relate γ_1 to v_1 and γ_2 to v_2 into Eqn.(4.12) gives the relative contamination ratio

$$\begin{aligned} |RC_{\gamma_1\gamma_2}| &= \frac{\exp^{-\Im(v_1)}[(\exp^{-2\Im(\Delta v)} - 2\exp^{-\Im(\Delta v)}\cos(\Re(\Delta v)) + 1]^{1/2}\exp^{\Im(v_1)}}{\exp^{\Im(v_1)}[(\exp^{+2\Im(\Delta v)} - 2\exp^{+\Im(\Delta v)}\cos(\Re(\Delta v)) + 1]^{1/2}\exp^{-\Im(v_1)}} \\ &= \frac{1}{\exp^{+\Im(\Delta v)}}. \end{aligned} \quad (4.17)$$

When $\Im(\Delta v)$ is positive, $|RC_{\gamma_1\gamma_2}|$ is less than one and the relative perturbation of γ_2 by the contamination is therefore greater than the relative perturbation of γ_1. However when $\Im(\Delta v)$ is negative $|RC_{\gamma_1\gamma_2}|$ is greater than one and the relative perturbation of γ_1 is greater than that in γ_2. In practice the sign of $\Im(\Delta v)$ changes in a random manner over the entire Fourier plane when an image is contaminated, consequently it is not possible to predict whether zero samples with smaller $|\gamma|$ are perturbed in relative terms less by the contamination than those of large $|\gamma|$ or vice versa. The above result has shown that the original belief that selecting zeros of small $|\gamma|$ could result in superior reconstructions is incorrect, since it appears, for wide-band contamination at least, that the relative perturbation on a zero does not decrease

with decreasing $|\gamma|$. As it is impossible to predict which zeros are perturbed most by the contamination in relative terms, any zero sample from the zero sheet is as good as any other for the purposes of image reconstruction. Experience with reconstructions from a variety of contaminated images lends support to this conclusion.

4.4.2 Band-limited Contamination

In the situation where the spectrum of an image's contamination occupies a region of Fourier space which is smaller than that occupied by the image's spectrum (i.e. the contamination is band-limited), it is reasonable to expect that the perturbation of the image's Z spectrum and zero sheet, resulting from this contamination, should also be confined to a finite region of Z space. This suggests the possibility, given prior knowledge of the contamination bandwidth, of developing an algorithm that selects zero samples only from the unperturbed regions of Z space. This in turn would allow an optimum reconstruction of the image to be obtained via the algebraic reconstruction technique described in §4.5.2. In this section the regions of an image's zero sheet that are not perturbed by three different forms of additive contamination are determined. The three types of contamination considered here are: high frequency contamination where all the image's frequency components above a particular frequency are contaminated, low frequency contamination where all the image's frequency components below a particular frequency in Fourier space are contaminated, and intermediate frequency contamination where an intermediate region of the image's Fourier spectrum is contaminated. In all the following derivations it is assumed that the cut-off frequencies of the contamination in Fourier space are the same in both the u and v directions. The lowest frequency component of the image corrupted by the contamination is denoted by the term BW_c^l and the highest frequency component is denoted by BW_c^h. It is also assumed, for simplicity, that the contamination has no frequency components above $u = \frac{1}{2\epsilon}$ and $v = \frac{1}{2\epsilon}$ and consequently, there is no need to be concerned with the affect of aliasing on the contaminated spectrum $\tilde{A}(u, v)$.

To determine what effect high frequency contamination has on an image's zero sheet, consider the case where a 2-D image $a(x, y)$ is corrupted by contamination that only distorts the frequency components of $A(u, v)$ which lie above the frequency BW_c^l. The extent of this high frequency contamination in analytically continued Fourier space is shown in Fig. 4.10 (a) and (c) where it can be seen that the contamination is confined to the region in Fourier space where $|\Re(u)|$ or $|\Re(v)|$ are greater than BW_c^l. It should be noted that the spectra in Fig. 4.12 are periodic, with period $\frac{1}{\epsilon}$, for the reasons outlined in §2.7. On using the mappings defined by Eqns. (2.70) and (2.71), which describe how points in Fourier space map to points in Z space, it can be seen (Fig. 4.10 (b) and (d)) that the contaminated region of Fourier space maps to the region in Z space where $|Ph(\zeta)| > \Omega BW_c^l$ or $|Ph(\gamma)| > \Omega BW_c^l$ ($\Omega = 2\pi\epsilon$).

The reasoning above predicts that zeros on a zero sheet slice having $|Ph(\gamma)| < \Omega BW_c^l$ and $|\phi| < \Omega BW_c^l$ are least likely to be affected by high frequency contamination. In Fig. 4.11 an ensemble of ten slices (ρ = 1.0) produced from an 8×8 random image corrupted with ten different equal energy versions of high frequency contamination with $BW_c^l = \frac{1}{4\epsilon}$ is shown ($BW_c^l = \frac{1}{4\epsilon}$ corresponds to the complete left hand side of the γ and ζ planes, i.e. γ or $\zeta > \frac{\pi}{2}$). On comparing individual zeros of the same $|\gamma|$ in Fig. 4.11, it can be seen that as $|Ph(\gamma)|$ decreases from π to 0 the lines representing the ensemble of tracks move closer together. This indicates that as $|Ph(\gamma)|$ tends to zero the tracks are perturbed less by the high frequency contamination, confirming the above prediction. More explicitly it can be seen in Fig. 4.11 that the zeros in the region marked A, corresponding to $|Ph(\gamma)| > \frac{\pi}{4}$, are perturbed more by the high frequency contamination than those in the regions marked B, corresponding

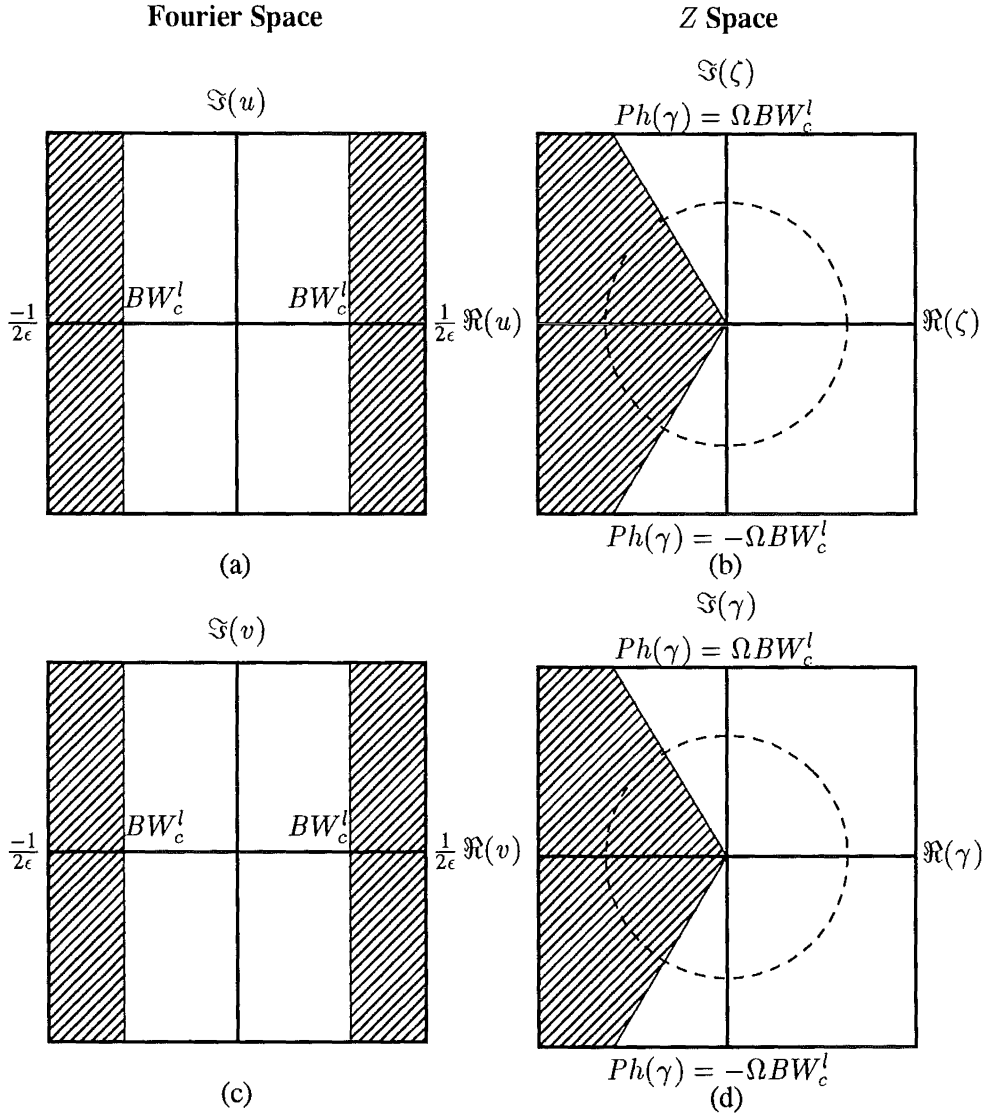

Figure 4.10 High frequency contamination in Fourier space of bandwidth BW_c^l in (a) the u plane and (c) the v plane. Also shown is the contamination in (a) and (b) mapped into the (c) ζ plane and (d) γ planes in Z space respectively.

to $|Ph(\gamma)| < \frac{\pi}{4}$. It is necessary to compare zeros of the same magnitude because as was described in §4.4.1, for a particular level of contamination, the perturbation of the zeros tends to increase in absolute terms with increasing magnitude of the zeros. As a general observation it can be seen that the contamination does perturb the zero tracks in the left hand side of the γ plane much more than those in the right hand side of the γ plane. The zeros for $\frac{\pi}{2} \leq |\phi| \leq \pi$ also appear to be perturbed more by the contamination than those for $|\phi| < \frac{\pi}{2}$ as predicted. The interwoven nature of the tracks, however, makes it extremely difficult to confirm this property and so the ϕ dependence has been omitted from Fig. 4.11 for reasons of clarity.

The set of figures and theoretical results presented above appear to indicate that if an image is corrupted with high frequency contamination and is to be reconstructed from its zero samples using the algebraic reconstruction technique (see §4.5.2), then it may be advantageous to select zeros mainly from those regions of Z space where $|Ph(\gamma)| < \Omega BW_c^l$ and $|\phi| < \Omega BW_c^l$, since zeros in these regions appear to be affected least by the contamination. The success

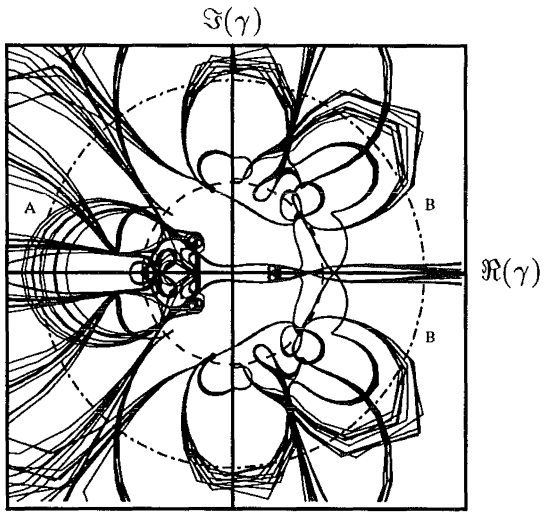

Figure 4.11 An ensemble of ten $\rho = 1.0$ slices for ten contaminated versions of the same 8×8 random image. The contamination was taken from a high frequency contamination distribution with $BW_c^l = \frac{1}{4\epsilon}$. Two circular contours are also shown using broken lines so that zeros of the same magnitude can be compared.

of this approach would naturally be dependent on obtaining enough samples of an accuracy sufficient to produce the required number of linearly independent equations required by the reconstruction technique.

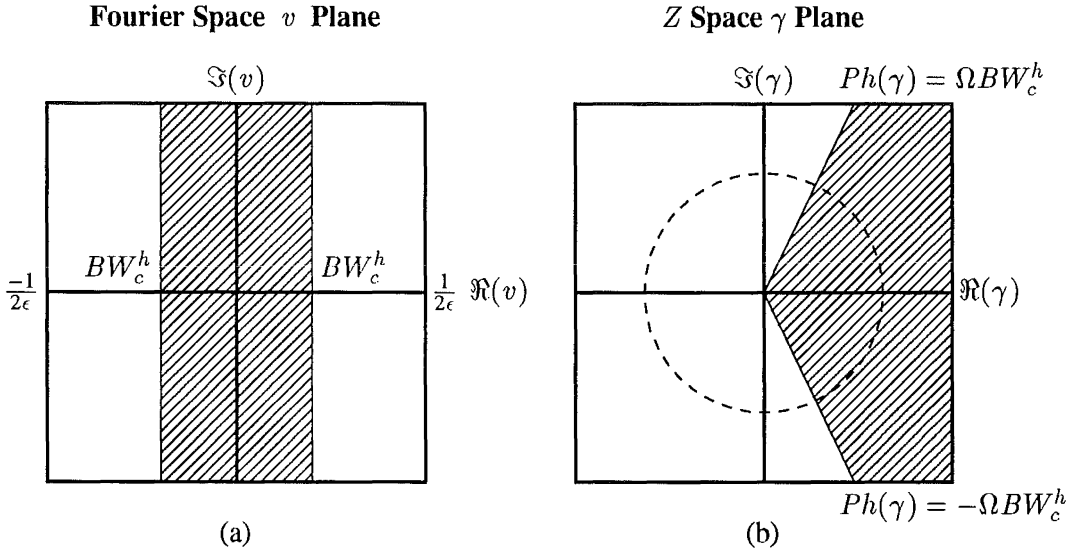

Figure 4.12 a) Low frequency contamination in Fourier space of bandwidth BW_c^h and b) the same contamination as in a) mapped into Z space.

For low frequency contamination, the Fourier/Z space mappings indicate that the region of Z space most likely to be affected is that where both $|Ph(\zeta)| < \Omega BW_c^h$ and $|Ph(\gamma)| < \Omega BW_c^h$. The region of Z space where the absolute phase of either of the Z space variables is greater than ΩBW_c^h is the region least likely to be affected by the contamination. This mapping of the low frequency contamination between Fourier space and Z space is shown pictorially in Fig. 4.12. An ensemble of ten slices (ρ=1.0) from the same 8×8 image used to produce Fig. 4.11, corrupted with ten different equal energy versions of low frequency contamination with $BW_c^h = \frac{1}{4\epsilon}$, is shown in Fig. 4.13. The contamination has thus been confined to the right hand

side of the ζ and γ planes. On comparing zeros of the same magnitude, again it can be seen that those zero tracks within the left hand side of γ space (i.e. those outside the bandwidth of the contamination) tend to be perturbed less than those in the right hand side of γ space. For example the zeros in the region labelled C and D in Fig. 4.13 tend to be perturbed less than those in the regions marked E and F respectively. The above results suggest that an optimal reconstruction of an image corrupted with low frequency contamination could be obtained from its zero sheet by using the fact that both $|\gamma|$ and $|\phi|$ must be less than ΩBW_c^h for a zero to lie within the contaminated region of Z space. This in turn suggests that, zeros could be selected from the region of Z space corresponding to $|Ph(\gamma)| > \Omega BW_c^h$, allowing the ϕ dependence to be ignored.

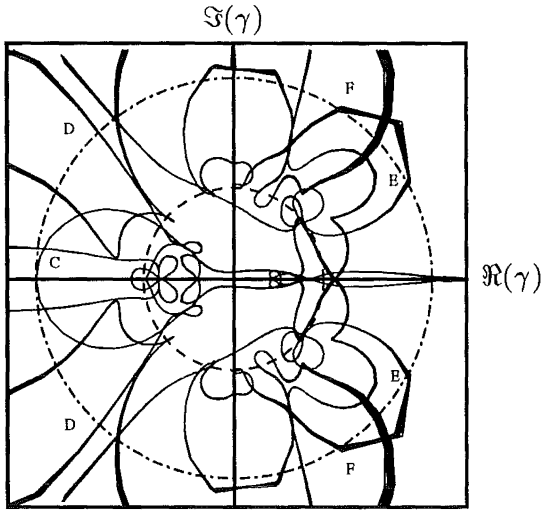

Figure 4.13 An ensemble of ten $\rho = 1.0$ slices for ten contaminated versions of the same 8×8 random image. The contamination was taken from a low frequency contamination distribution with $BW_c^h = \frac{1}{4\epsilon}$. Two circular contours are also shown using broken lines so that zeros of the same magnitude can be compared.

It is also interesting to compare the ensembles from the high and low frequency cases as it can be seen that tracks that were significantly perturbed in Fig. 4.11 (i.e. those in the region marked A) are not perturbed in Fig. 4.13 (i.e. those in the region marked C). On comparing the two ensembles it should also be evident that the overall perturbation of the slices is more marked in the high frequency situation than the low frequency situation, even though both sets of images have the same SNRs. The reason for this is that the image used here is typical of most images encountered in practice which have a power spectrum that decreases rather rapidly with increasing frequency. The high frequency contamination therefore perturbs the high frequency components of the image in relative terms more than low frequency contamination can perturb the low frequency components of the image and as a result, the zeros at high frequencies are perturbed more than those at low frequencies. Results described in §5.3.3 confirm that this is indeed the case.

The final form of band-limited contamination to be described here is intermediate frequency contamination for which $BW_c^l < BW_c^h$ (see Fig. 4.14 (a)). Consequently, in Z space the uncontaminated regions are those where $|Ph(\gamma)|$ or $|Ph(\zeta)|$ are greater than ΩBW_c^h or where both $|Ph(\gamma)|$ and $|Ph(\zeta)|$ are less than ΩBW_c^l as is shown in Fig. 4.14 (b). As with the previous two types of contamination, an ensemble containing ten superimposed zero slices (ρ = 1.0) from a 8×8 random image contaminated with ten different versions of band-limited contamination of the same energy was produced. In this case the ten different versions of

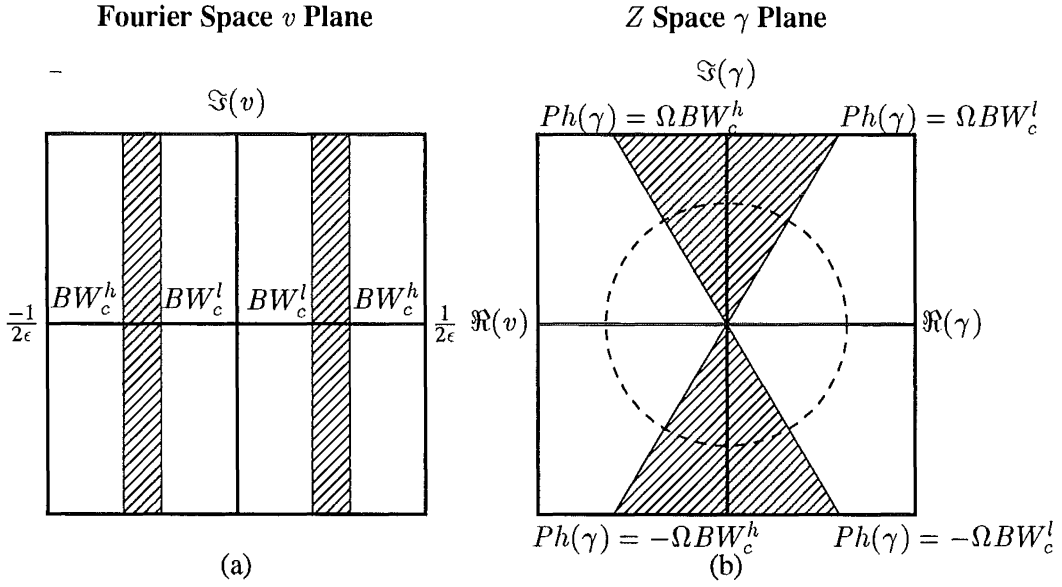

Figure 4.14 a) Intermediate frequency contamination in Fourier space which contaminates frequencies between BW_c^l and BW_c^l. b) The same contamination as in a) mapped into Z space.

band-limited frequency contamination was confined to the frequencies bounded by $BW_c^l = \frac{1}{8\epsilon}$ and $BW_c^h = \frac{3}{8\epsilon}$ or equivalently the regions of Z space where $\frac{\pi}{4} < |Ph(\zeta)| < \frac{3\pi}{4}$ and $\frac{\pi}{4} < |Ph(\gamma)| < \frac{3\pi}{4}$. The resulting ensemble is shown in Fig. 4.15 and on comparing Fig. 4.15 to Figs. 4.11 and 4.13 the fact that the contamination is confined to the above regions of Z space can most clearly be seen. For example, the low level of very high frequency contamination can be seen if the tracks at the point marked G in Fig. 4.15, A in Fig. 4.11, and C in Fig. 4.13 are compared. The tracks in Fig. 4.15 and Fig. 4.13 are almost exactly overlaid, indicating a low level of distortion, while in the high frequency contamination case (Fig. 4.11) these tracks are quite widely spread. On comparing the very low frequency region of Fig. 4.15 at the point marked H to the corresponding points in Figs. 4.11 and 4.13 the amount of perturbation of these tracks is marginally more than that in Figs. 4.11 and marginally less than that in Fig. 4.13. Finally, the amount of zero track perturbation in the region marked I in Fig. 4.15 (i.e. $\frac{\pi}{4} < |Ph(\gamma)| < \frac{3\pi}{4}$) is considerably greater than that in the corresponding regions in Figs. 4.11 and 4.13. Thus, it appears that in the intermediate frequency contamination case in Fig. 4.15 the contamination is indeed predominantly confined to the region $\frac{\pi}{4} < |Ph(\gamma)| < \frac{3\pi}{4}$.

4.4.3 Contaminated Convolutions and Power Spectra

The effects of contamination on the zero sheets of convolutions and power spectra are considered in this section. The contamination is found to distort the two component zero sheets in a manner which prevents the elementary deconvolution and phase retrieval algorithms of Lane *et al.* (1987), as described in §4.3, from functioning. As it is intended that the new zero-based algorithms described in §5.1 and §6.1.2 should be able to deblur or retrieve the phase of contaminated images it is necessary to describe the properties of the zero sheets belonging to a reducible image after contamination is added. In this section only the properties of a zero sheet of a blurred image corrupted with additive contamination are described, however, the properties described here apply equally well to any form of reducible image.

In the situation where a blurred and contaminated image can be described by Eqn.(2.49)

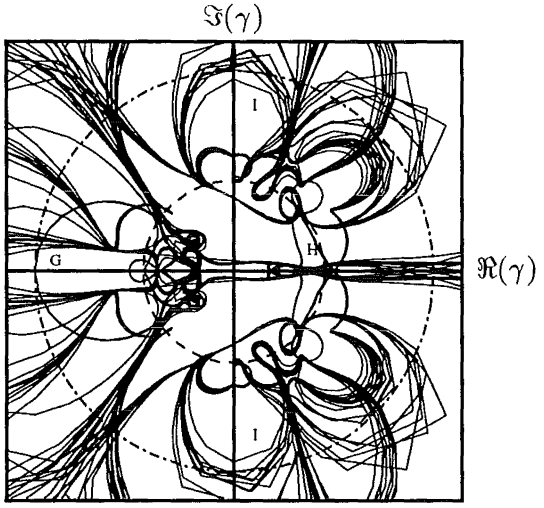

Figure 4.15 An ensemble of ten $\rho = 1.0$ slices for ten contaminated versions of the same 8×8 random image. The contamination was taken from a intermediate frequency contamination distribution with $BW_c^l = \frac{1}{4\epsilon}$ and $BW_c^h = \frac{3}{4\epsilon}$. Two circular contours are also shown using broken lines so that zeros of the same magnitude can be compared.

the Z spectrum of this image can be expressed in the form

$$\mathcal{G}(\zeta,\gamma) = \mathcal{F}(\zeta,\gamma)\mathcal{H}(\zeta,\gamma) + \mathcal{C}(\zeta,\gamma), \tag{4.18}$$

where $\mathcal{F}(\zeta,\gamma)$, $\mathcal{H}(\zeta,\gamma)$, and $\mathcal{C}(\zeta,\gamma)$ are the Z spectra of the original image, psf, and contamination respectively. The contamination term in the above expression for $\mathcal{G}(\zeta,\gamma)$ has the effect of turning the polynomial describing the spectrum of the blurred image from a reducible polynomial, with two factors $\mathcal{F}(\zeta,\gamma)$ and $\mathcal{H}(\zeta,\gamma)$, into an irreducible polynomial which cannot be factorised. This in turn implies that the set of zeros belonging to $\mathcal{G}(\zeta,\gamma)$ cannot be automatically separated into the two individual sets of zeros which characterise the blurred image's components and form their zero sheets, i.e.

$$\mathrm{Z}\{\mathcal{G}(\zeta,\gamma)\} \neq \mathrm{Z}\{\mathcal{F}(\zeta,\gamma)\} \cup \mathrm{Z}\{\mathcal{H}(\zeta,\gamma)\}.$$

Thus, the zero sheet of $\mathcal{G}(\zeta,\gamma)$ is a single 2-D surface in 4-D space and the algorithms described in §4.3 can no longer be used to automatically separate the zero sheet of the blurred image into the two separate surfaces which represent $\mathcal{F}(\zeta,\gamma)$ and $\mathcal{H}(\zeta,\gamma)$.

In order to establish how, in the presence of contamination, the two separate zero sheets pertaining to the component images become a single zero sheet, the analytic nature of the surfaces must be considered. It was described and also shown in §4.4.1 that when wide-band contamination is added to an image, the entire zero sheet of the image is perturbed in a smooth continuous manner, with the size of the perturbations being dependent on the energy of the contamination. Therefore if an infinitesimal amount of contamination is added to a blurred image, the two separate surfaces representing the zero sheets of the individual components must merge together at the regions in 4-D space where the two separate zero sheets are closest (i.e. at the points where the two surfaces intersect) to form the single zero sheet. The fact that the two component zero sheets merge together at the localised regions in Z space where the two surfaces are in close proximity was first noted and shown by Lane and Bates (1987). Another interesting property is that as the level of contamination increases and the individual surfaces become more distorted the number of regions in Z space where the two surfaces have merged together in Z space also increases (Lane and Bates, 1987). This is because the

increasing contamination increases the distortion of the zero sheets, which in turn tends to produce an increasing number of regions in Z space where the two surfaces are close enough together to merge (see Fig. 4.16). As the level of contamination increases it is also observed that the regions on the single zero sheet where the two separate surfaces are interconnected gradually increase in size due to the increasing distortion of the zero sheet as a whole (Lane and Bates, 1987). In addition to the above, it has been the author's experience that the merging of the zero sheets becomes less frequent as ρ tends to either zero or infinity. This is because the zero sheets contract to point zeros as ρ tends to zero or infinity (see §4.1.1), which decreases the likelihood of the two sheets being close enough together for merging to occur in these regions of Z space. In Fig. 4.16 a series of $\rho = 0.7$ slices for the zero sheet belonging to the convolution initially displayed in Fig. 4.4 are shown for increasing levels of contamination. It can be seen in this set of figures, where the regions of merging are inside the superimposed circles, that: the zero sheets do indeed merge where they are closest together, there is an increase in the predominance of merging as the contamination increases, and as the contamination increases the size of the interconnecting regions between the two sheets also increases.

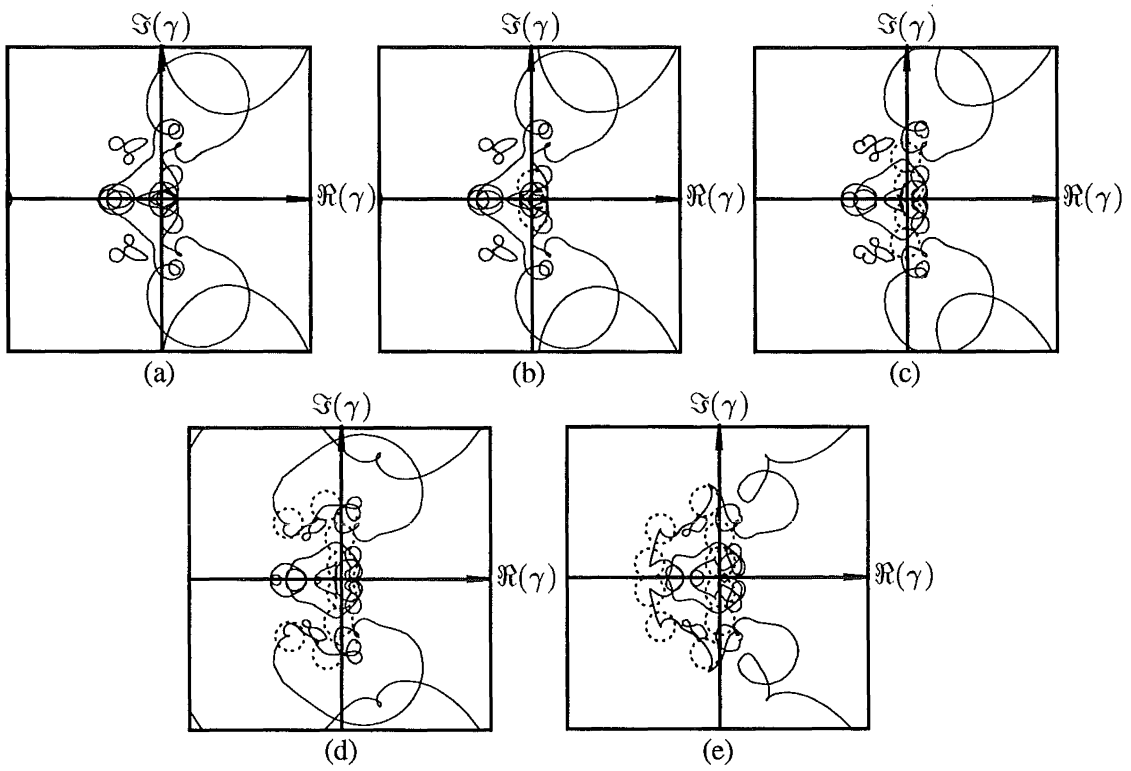

Figure 4.16 The merging together of the component zero sheets of a convolution as the SNR of the convolution decreases. Shown is the zero sheet slice at $\rho = 0.8$ for SNRs of (a) 60 dB, (b) 50 dB, (c) 40 dB, and (d) 30 dB. The regions were merging has occurred are indicated by a dotted circle.

The fact that for modest contamination levels the merging is localised and the zero sheets are only slightly perturbed by the contamination suggests that it should be possible to recognise these localised regions where the sheets are merged and the sheets which pertain to the individual components of the convolution (see Fig. 4.17 (a))(Lane and Bates, 1987). For modest contamination levels is is therefore possible to express the zero sheet of the blurred image in the form

$$\mathrm{Z}\{\mathcal{G}(\zeta,\gamma)\} \approx \mathrm{Z}\{\mathcal{F}(\zeta,\gamma)\} \cup \mathrm{Z}\{\mathcal{H}(\zeta,\gamma)\}, \tag{4.19}$$

and it should be possible to separate the two component sheets allowing an estimate of either

$f(x, y)$ or $h(x, y)$ to be obtained. In Fig. 4.17 four sets of $\rho = 0.6$ superimposed slices for the convolution originally depicted in Fig. 4.4 are presented. Each of these sets of slices were produced by adding ten different samples of wide-band contamination, with the specified amplitude, to the original convolution to produce four ensembles with SNRs of 60 dB, 50 dB, 40 dB, and 30 dB. The set of slices in Fig. 4.17 demonstrate that the union relationship in Eqn.(4.19) should hold for blurred image SNRs of greater than approximately 30 dB. As the SNR worsens towards 30 dB it can be seen in Fig. 4.19 (d) that the individual zero sheets become increasingly difficult to recognise and, since the sheets are severely distorted, merging no longer tends to be localised. At SNR's of 30 dB it is therefore probably not valid to assume that the zero sheet of a blurred image can be adequately described by Eqn.(4.19).

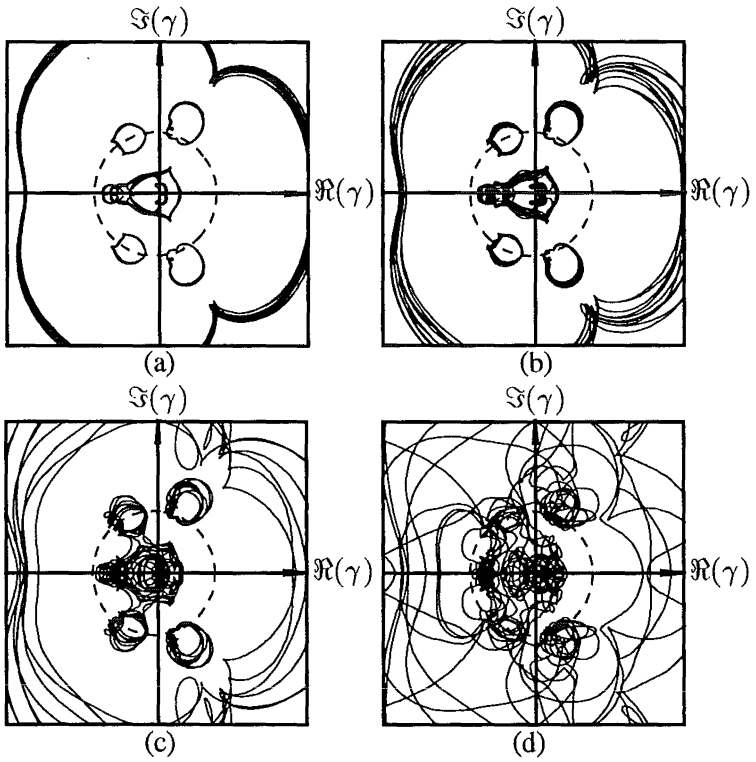

Figure 4.17 The merging together of the component convolution zero sheets as the SNR decreases which makes recognition of the individual components difficult. Shown are ensembles produced from the same convolution corrupted with contamination to give the convolution a (a) 60 dB, (b) 50 dB, (c) 40 dB, and (d) 30 dB SNR.

4.5 IMAGE RECONSTRUCTION FROM ZERO SHEETS

4.5.1 Fourier Reconstruction

Once an approximation to an image's zero sheet has been found, for example in the case where zero sheet separation is being used to solve the image blurring or phase retrieval problem, a technique which allows an estimate of an image to be reconstructed from samples of its zero sheet is required. In this section the Fourier reconstruction technique is described. The principle of the method is to estimate (from Z space zero samples) the Fourier transform of the image on the real (u, v) plane at the points necessary to allow the IDFT to be used to reconstruct an image estimate (Lane and Bates, 1987; Lane *et al.*, 1987; Parker, 1994; Satherley, 1994).

The Fourier reconstruction technique is based around Eqn.(2.70) and Eqn.(2.71) which describe how points in Z space map to corresponding points in Fourier space. This mapping was first introduced in §2.7 and as alluded to in §2.7, points in Z space with unit magnitude map to points in Fourier space on the real u, v plane. Consider the zero sheet $\mathrm{Z}\{\hat{A}(\zeta, \gamma)\}$ which is an estimate of the zero sheet that belongs to the $N \times N$ pixellated image $a(x, y)$. A consequence of the above mappings is that the $N-1$ point zeros $\hat{\gamma}_j(\zeta_c)$; $j = 1 \cdots, N-1$, produced by setting $\zeta = \zeta_c = 1.0 \exp^{i2\pi u_c \epsilon}$ (see Eqn.(4.7)), where u_c is real, characterise the 1-D polynomial $\hat{A}(\zeta_c, \gamma)$ which maps to the line $u = u_c$ on the real u, v plane in Fourier space. If γ in this 1-D polynomial is now set to $\gamma_c = 1.0 \exp^{i2\pi v_c \epsilon}$, where v_c is real, then $\hat{A}(\zeta_c, \gamma_c)$ provides the value of the Fourier transform of the image estimate, at the position (u_c, v_c), to within an arbitrary complex constant. Thus, by setting ζ and γ in the above manner it is possible to reconstruct an estimate of the Fourier spectrum on the real u, v plane to within a complex constant. Determining the constant is discussed below. It should be noted that the choice of the roles of ζ and γ in the above discussion is arbitrary and the roles can be reversed if desired.

To reconstruct an estimate of $a(x, y)$ by using the DFT, an estimate of $A(u, v)$ must be known at least at the equispaced points (u_n, v_m), where $n = 0, 1, \cdots, N-1$, and $m = 0, 1, \cdots, N-1$. Obtaining samples of $\hat{A}(u, v)$ in the above manner ensures that the sampling theorem is satisfied (see §2.5.4). Thus, the Fourier reconstruction technique requires that the set of polynomials $\hat{A}(\zeta_n, \gamma)$, where $\zeta_n = 1.0 \exp^{i2\pi n/N}$ and $n = 0, 1, \cdots, N-1$, are formed. The $\hat{A}(\zeta_n, \gamma)$ are polynomials of the form

$$\hat{A}(\zeta_n, \gamma) = WC(\zeta_n) \prod_{j=1}^{N-1} (\gamma - \hat{\gamma}_j(\zeta_n)), \tag{4.20}$$

where $C(\zeta_n)$ is a complex valued scale factor that is different for each of the 1-D polynomials, and W is a complex constant. To obtain the values of $\hat{A}(u_n, v_m)$ required to allow the DFT to be used to reconstruct the image, each of the 1-D polynomials, in the form of Eqn.(4.20), is evaluated for $\gamma = \gamma_m = 1.0 \exp^{i2\pi m/N}$, where $m = 0, 1, \cdots, N-1$. A pictorial demonstration of how the Fourier spectrum is reconstructed along the line $u = u_1$ is shown in Fig. 4.18 for a 4×4 image. In Fig. 4.18 (a) the four preset values of ζ required to reconstruct the four 1-D polynomials that map to the lines $u = u_n$ in Fourier space are shown. Depicted in Fig. 4.18 (b) are the four point zeros, produced by setting $\zeta = \zeta_1$, that give rise to the 1-D polynomial

$$\hat{A}(\zeta_1, \gamma) = (\gamma - \hat{\gamma}_1(\zeta_1))(\gamma - \hat{\gamma}_2(\zeta_1))(\gamma - \hat{\gamma}_3(\zeta_1))(\gamma - \hat{\gamma}_4(\zeta_1)), \tag{4.21}$$

which maps to the line $u = u_1$ in Fourier space. On evaluating Eqn.(4.21) at $\gamma = 1.0 \exp^{i2\pi v \epsilon}$, for all real values of v, the function $\hat{A}(u_1, v)$ describing the spectrum along the line u_1 in Fig. 4.18 (c) can be produced. Similarly, be reversing the roles of ζ and γ the function $\hat{A}(u, v_1)$ in 4.18 (c) can be produced.

Before the 1-D polynomials are evaluated to find the $\hat{A}(u_n, v_m)$, it is first necessary to remove the $C(\zeta_n)$ that scale each of the 1-D polynomials. Once this has been done $\hat{A}(u, v)$ can be calculated at the necessary sample points to within the single unresolvable complex constant W. To remove the $C(\zeta_n)$ a 1-D polynomial $\hat{A}(\zeta, \gamma_m)$ that maps to the line $v = v_m$, and as a consequence describes the spectrum along a single line perpendicular to the lines which the $\hat{A}(\zeta_n, \gamma)$ in Eqn.(4.20) map to, is produced. This 1-D polynomial

$$\hat{A}(\zeta, \gamma_m) = W(\gamma_m) \prod_{j=1}^{N-1} (\zeta - \hat{\zeta}_j(\gamma_m)), \tag{4.22}$$

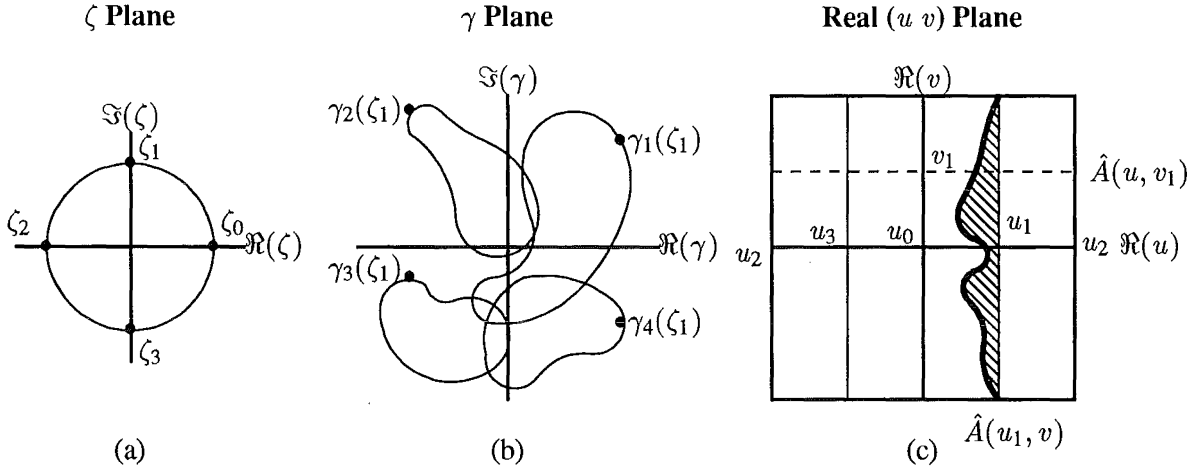

Figure 4.18 Shown in (c) are the lines in Fourier space along which $\hat{A}(u, v)$ must be estimated to reconstruct a 4×4 image $\hat{a}(x, y)$ via the IDFT. Shown is the case where (a) $\zeta = \zeta_1$ and (b) the resulting zeros in the γ plane $\hat{\gamma}_j(\zeta_1)$ are used to reconstruct the polynomial $\hat{A}(u_1, v)$ that estimates $A(u, v)$ along the line $u = u_1$ in Fourier space.

is produced from the zeros $\hat{\zeta}_j(\gamma_m)$ that correspond to setting $\gamma_m = 1.0 \exp^{i2\pi m/N}$, where m is chosen from the range 0 to $N - 1$. In Eqn.(4.22) $W(\gamma_m)$ is equal to the complex constant W in Eqn.(4.20) and describes the unknown complex scaling factor that the Fourier spectrum, calculated via $\hat{A}(\zeta, \gamma_m)$, differs from $\hat{A}(u, v_m)$ by. At (ζ_n, γ_m) the values of $\hat{A}(\zeta, \gamma)$ provided by the 1-D polynomials defined by Eqn.(4.20) and the 1-D polynomial defined by Eqn.(4.22) must be equal. Thus, the $C(\zeta_n)$ can be calculated by equating Eqn.(4.20) and Eqn.(4.22) at each value of ζ_n to give

$$C(\zeta_n) = \frac{\prod_{j=1}^{N-1}(\zeta_n - \hat{\zeta}_j(\gamma_m))}{\prod_{j=1}^{N-1}(\gamma_m - \hat{\gamma}_j(\zeta_n))}, \qquad \forall n. \tag{4.23}$$

After dividing the $\hat{A}(\zeta_n, \gamma)$ by the corresponding $C(\zeta_n)$ and evaluating the scaled results at the Fourier sample points (as described above) an estimate of $a(x, y)$ can be obtained by applying the IDFT. The estimate of $a(x, y)$ differs from $a(x, y)$ by the unresolvable complex constant W, the inaccuracies in the $\hat{\zeta}_j(\gamma_m)$ and $\hat{\gamma}_j(\zeta_n)$, and the computational inaccuracies.

The accuracy of the normalisation process, designed to remove the $C(\zeta_n)$, is heavily dependent on the accuracy of the $\hat{\zeta}_j(\gamma_m)$ and the resulting 1-D polynomial $\hat{A}(\zeta, \gamma_m)$. If $\hat{A}(\zeta, \gamma_m)$ is erroneous, it is clear that the $\hat{A}(\zeta_n, \gamma)$ will be scaled incorrectly with the resulting $\hat{a}(x, y)$ also being in error. Using the single 1-D polynomial $\hat{A}(\zeta, \gamma_m)$ to perform the normalisation is clearly a weakness of the reconstruction technique described above. It should be noted that the Fourier reconstruction process used to produce the results in §5.5 performed the normalisation with a single 1-D polynomial. Recently, however, Ghiglia *et al.* (1993) have reported a new more robust normalisation technique. The technique of Ghiglia *et al.* (1993) reconstructs the spectral samples along the columns using the technique described above, then reconstructs the spectral samples across the rows (not a single row as described above), and normalises all the rows and columns by using a least squares technique. Explicitly, Ghiglia *et al.* (1993) compute the matrices

$$\tilde{\mathbf{A}} = \mathbf{cA} \tag{4.24}$$

and

$$\check{\mathbf{A}} = \mathbf{r}\mathbf{A} \tag{4.25}$$

where $\mathbf{A}$ is the $N \times N$ matrix containing the Fourier samples to within the single unresolvable complex constant, $\mathbf{c}$ is a vector containing the complex scale factors for each column, $\mathbf{r}$ is a vector containing the complex scale factors for each row, $\tilde{\mathbf{A}}$ is the matrix of Fourier samples produced column by column, and $\check{\mathbf{A}}$ is the matrix of Fourier samples produced row by row. On equating Eqn.(4.24) to Eqn.(4.25) the set of equations

$$\mathbf{r}\tilde{\mathbf{A}} - \mathbf{c}\check{\mathbf{A}} = 0 \tag{4.26}$$

which must be solved to find the $\mathbf{r}_m$ and $\mathbf{c}_n$ is obtained. Eqn.(4.26) can be written in the form of the matrix equation

$$\boldsymbol{\Gamma}\mathbf{d} = 0, \tag{4.27}$$

where $\mathbf{d} = [\mathbf{r}_1, \cdots, \mathbf{r}_{N-1}, \mathbf{c}_1, \cdots, \mathbf{c}_{N-1}]^T$ and $\boldsymbol{\Gamma}$ is a sparse matrix containing the $\tilde{\mathbf{A}}_{n,m}$ and $\check{\mathbf{A}}_{n,m}$ (see Ghiglia *et al.* (1993)). To solve Eqn.(4.27) in a least squares manner for $\mathbf{d}$, the non-zero solutions to the equation

$$\boldsymbol{\Gamma}^T\boldsymbol{\Gamma}\mathbf{d} = 0, \tag{4.28}$$

are found (see §4.5.2.2 for the matrix least squares principle). In fact Ghiglia *et al.* (1993) solve Eqn.(4.28) using an eigenvector/eigenvalue approach because they use the size of the smallest eigenvalue to determine whether $\tilde{\mathbf{A}}$ and $\check{\mathbf{A}}$ do indeed pertain to the same factor in a convolution.

It is indicated in §4.5.3 that one of the weaknesses of the Fourier reconstruction technique is that it demands that a specific small set of zeros samples, which may be contaminated more than other zero samples, be selected from the zero sheet. As reported in §4.5.3, Satherley and Parker (1993) have recently developed an adapted Fourier reconstruction technique that offers some flexibility over which zero samples must be selected. The principle of the algorithm of Satherley and Parker (1993) is that an estimate of the Fourier spectrum of the image is reconstructed on a complex plane in Fourier space (i.e. not the real u, v plane as described above). A simple scaling operation in image space then allows the image reconstructed from the complex Fourier plane to be mapped to that which would have been produced from the spectrum on the real u, v plane. More explicitly, Satherley and Parker (1993) use the zeros produced by setting $\zeta_n = \exp^{i2\pi(n+ip)/N}$, where $n = 0, 1, \cdots, N-1$ and p is a real constant, to produce the set of 1-D polynomials $\tilde{A}(\zeta_n, \gamma)$ that map onto the line $u = u_n + ip$ in Fourier space (c.f. the standard method described earlier). By evaluating the $\tilde{A}(\zeta_n, \gamma)$ at $\gamma_m = \exp^{i2\pi(m+iq)/N}$, where $m = 0, 1, \cdots, N-1$ and q is a real constant, the Fourier transform samples $\tilde{A}(u_n + ia, v_m + ib)$ on a complex Fourier plane are obtained. On examining the modulation property of the Fourier transform in Table 2.1, it becomes clear how the image obtained from $\tilde{A}(u_n + ia, v_m + ib)$, via the IDFT, can be scaled to obtain $\hat{a}(x, y)$. Specifically, the modulation property indicates that

$$\hat{a}(x, y) = \frac{\text{IDFT}\{\tilde{A}(u_n + ia, v_m + ib)\}}{\exp^{-2\pi(pn+qm)/N}}, \tag{4.29}$$

where $n = 0, 1, \cdots, N-1$ and $m = 0, 1, \cdots, N-1$, gives the required image estimate. Satherley *et al.* (1994) note that the accuracy of the image reconstructed using this technique tends to deteriorate in the regions corresponding to a small denominator in Eqn.(4.29). They have shown that this problem can be alleviated considerably by reconstructing a number of $\hat{a}(x, y)$, from various values of p and q, and combining the resulting reconstructions.

4.5.2 Algebraic Reconstruction

4.5.2.1 The Principle of Algebraic Reconstruction

The second method which can be used to reconstruct an image from samples of its zero sheet is the algebraic reconstruction technique. The algebraic reconstruction technique obtains a reconstruction by first constructing an overdetermined system of linear equations, from a set of zero samples, in which the unknowns are the pixel values of the image (Izraelevitz and Lim, 1987). The resulting system can then be solved to obtain an estimate of the image.

To construct the system of linear equations a set of samples from the image's zero sheet, denoted here by $(\zeta_{s\lambda}, \gamma_{s\lambda})$, $\lambda = 1, \cdots, \Lambda$, are selected in some manner. Each of the $(\zeta_{s\lambda}, \gamma_{s\lambda})$ are then substituted into the 2-D Z transform (Eqn.(2.69)) for the pixellated $N \times N$ image to obtain a set of equations of the form

$$\begin{bmatrix} \zeta_{s1} & \gamma_{s1} & \zeta_{s1}\gamma_{s1} & \cdots & \zeta_{s1}^{N}\gamma_{s1}^{N} \\ \zeta_{s2} & \gamma_{s2} & \zeta_{s2}\gamma_{s2} & \cdots & \zeta_{s2}^{N}\gamma_{s2}^{N} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \zeta_{s\Lambda} & \gamma_{s\Lambda} & \zeta_{s\Lambda}\gamma_{s\Lambda} & \cdots & \zeta_{s\Lambda}^{N}\gamma_{s\Lambda}^{N} \end{bmatrix} \begin{bmatrix} f_{01} \\ f_{10} \\ f_{11} \\ \vdots \\ f_{NN} \end{bmatrix} = - \begin{bmatrix} f_{00} \\ \vdots \\ f_{00} \end{bmatrix}, \tag{4.30}$$

where the pixel values f_{nm} are the unknowns. In order to be able to solve the above system of equations it is necessary to set f_{00} to unity to prevent the trivial solution (i.e. $f_{nm} = 0, \forall n$ and m) being obtained. In principle the zero samples that are required to construct the system of equations can be selected from anywhere on the zero sheet, however the finite resolution of the number representation scheme in the computer used and the inevitable zero sheet contamination dictate that the zero samples must be spread reasonably widely in Z space to ensure linear independence between the equations comprising Eqn.(4.30). Solving the system of equations gives an estimate of the image scaled by an arbitrary complex constant.

According to the reasoning given in §4.1.2, two distinct images can have at most $2(N-1)^2$ zeros in common, consequently $\Lambda \geq 2(N-1)^2 + 1$ to guarantee a unique solution. The system of equations, required to guarantee a unique solution, is therefore a system of $2(N-1)^2 + 1$ equations in $N^2 - 1$ unknowns. A solution can of course be obtained from any $N^2 - 1$ of these equations and Izraelevitz and Lim (1987) have found that if the zero samples are chosen at random then the $N^2 - 1$ zero samples are almost always sufficient to provide a unique solution. This is because the probability of selecting a complete set of zeros that are common to another image is very small, although this is of course conditional on the system of equations being of sufficient accuracy that numerical difficulties are avoided (Curtis and Oppenheim, 1987). The major advantage of using significantly more than the minimum required, is that an optimal solution in the least squares sense may be obtained in the presence of contamination. In theory Λ could be infinite, since the zero sheet is formed from an infinity of zeros, however the system of equations becomes increasingly ill-conditioned as Λ rises. The author uses the QR decomposition technique to solve the system of equations.

4.5.2.2 Obtaining an Optimal Solution (QR Decomposition)

The QR decomposition algorithm is a procedure which solves overdetermined systems of linear equations in an optimal manner by minimising the least squares error in the solution.

It functions by separating the matrix representing the system of equations into two separate matrices **Q** and **R**, which in turn allow a solution to be obtained by using the simple method of back substitution. In the situation where the matrix decomposition is performed by a series of Householder transformations, the QR decomposition process is both efficient and numerically stable (Strang, 1980). The advantages that QR decomposition has over simpler matrix inversion techniques are that all the equations in the system contribute to the solution and the accuracy of the solution is not as sensitive to the numerical problems associated with divide by zero errors. In this section the principles of matrix least squares, QR decomposition, and Householder transformations are described.

Consider an overdetermined system of equations of the form

$$\mathbf{Ax} = \mathbf{y}, \tag{4.31}$$

where the $M \times P$ matrix $\mathbf{A}$, $M > P$, is the matrix of coefficients, the vector $\mathbf{x}$ of extent P is the solution vector, and $\mathbf{y}$ is the vector of known constants which has an extent M. To solve the above system of equations, in a manner that permits all the information contained within the M equations comprising $\mathbf{A}$ to be used, a least squares technique can be employed. Least squares techniques aim to minimise the difference between the actual solution $\mathbf{x}$ and the solution obtained by the technique, i.e. the technique aims to minimise

$$\frac{1}{P}\sum_{i=1}^{P}[\mathbf{x}_i - \hat{\mathbf{x}}_i]^2 = \frac{1}{P}\sum_{i=1}^{P}[\mathbf{r}_i]^2 = E^2,$$

where $\hat{\mathbf{x}}$ is the solution obtained from the technique, the $\mathbf{r}_i$ are known as the residuals and E^2 is the mean square error. In terms of the matrix system of equations in Eqn.(4.31) the mean square error is given by the expression

$$\begin{aligned} E^2 &= [\mathbf{Ax} - \mathbf{y}]^T[\mathbf{Ax} - \mathbf{y}] \\ &= \mathbf{x}^T\mathbf{A}^T\mathbf{Ax} - \mathbf{x}^T\mathbf{A}^T\mathbf{y} - \mathbf{y}^T\mathbf{Ax} + \mathbf{y}^T\mathbf{y}. \end{aligned} \tag{4.32}$$

To determine the value of $\mathbf{x}$ which minimises E^2, Eqn.(4.32) is differentiated and equated to zero. Performing the differentiation gives

$$\frac{\delta E^2}{\delta \mathbf{x}} = 2\mathbf{x}^T\mathbf{A}^T\mathbf{A} - \mathbf{y}^T\mathbf{A} - \mathbf{y}^T\mathbf{A} + 0,$$

and on equating $\frac{\delta E^2}{\delta \mathbf{x}}$ to zero (Strang, 1980; Sinha and Kuszta, 1983)

$$\hat{\mathbf{x}} = [\mathbf{A}^T\mathbf{A}]^{-1}\mathbf{A}^T\mathbf{y}. \tag{4.33}$$

It should be noted that in the situation where some of the equations in $\mathbf{A}$ are known more accurately than others it is possible to weight the residuals so that the equations which are known most accurately contribute most to E^2. This weighting operation can be performed by using a system of equations of the form

$$\mathbf{Gr} = \mathbf{Gb} - \mathbf{GAx}, \tag{4.34}$$

where $\mathbf{G}$ is the $M \times M$ weighting matrix which only has nonzero elements in its leading diagonal. When the ith equation in $\mathbf{A}$ is known most accurately, the element $\mathbf{G_{ii}}$ is given the largest value in the matrix and the solution obtained therefore satisfies this most accurate equation to a larger degree than the other equations comprising $\mathbf{A}$ (Lawson and Hanson, 1974).

If the least squares solution was to be directly calculated using Eqn.(4.33) the solution would suffer from the numerical instability that commonly occurs when matrix inversions have to be performed. To avoid this instability it is possible to make a substitution, which is now described, that allows $\hat{\mathbf{x}}$ to be calculated via a simple back substitution. In the situation where $\mathbf{A}$ has linearly independent columns it is possible to factorise $\mathbf{A}$ into the form

$$\mathbf{A} = \mathbf{QR}, \tag{4.35}$$

where $\mathbf{Q}$ is a $M \times M$ orthonormal matrix (i.e. if $\mathbf{q}_{ci}$ is the ith column of $\mathbf{Q}$, then $\mathbf{q}_{ci}\mathbf{q}_{cj} = 0$ whenever $i \neq j$ and $\mathbf{q}_{ci}\mathbf{q}_{cj} = 1$ whenever $i = j$) and $\mathbf{R}$ is a $M \times P$ upper triangular matrix (Dongarra *et al.*, 1979; Strang, 1980). On substituting Eqn.(4.35) into Eqn.(4.33) it is possible to express $\hat{\mathbf{x}}$ in the form

$$\begin{aligned} \hat{\mathbf{x}} &= [\mathbf{R}^T\mathbf{Q}^T\mathbf{QR}]^{-1}\mathbf{R}^T\mathbf{Q}^T\mathbf{y} \\ &= [\mathbf{R}^T\mathbf{R}]^{-1}\mathbf{R}^T\mathbf{Q}^T\mathbf{y}, \end{aligned} \tag{4.36}$$

on using the relationship $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. Rearranging gives

$$\not{\mathbf{R}}^T\mathbf{R}\hat{\mathbf{x}} = \not{\mathbf{R}}^T\mathbf{Q}^T\mathbf{y}$$

and

$$\mathbf{R}\hat{\mathbf{x}} = \mathbf{Q}^T\mathbf{y}, \tag{4.37}$$

It can be seen from Eqn.(4.37) that once the QR decomposition has been performed, a simple back substitution operation allows $\hat{\mathbf{x}}$ to be determined. The most straightforward method available for performing the QR decomposition is the Gram-Schmidt process which takes a matrix with linearly independent columns $[\mathbf{a}_{c1}, \mathbf{a}_{c2}, \ldots, \mathbf{a}_{cP}]$, uses a series of projections to remove the components of each $\mathbf{a}_{cp}$ which are not aligned with a specific basis vector $\mathbf{q}_{cp}$, and thereby creates the orthonormal basis $[\mathbf{q}_{c1}, \mathbf{q}_{c2}, \ldots, \mathbf{q}_{cP}]$ for the space spanned by $\mathbf{A}$ (the $\mathbf{R}_p$ are also available directly from the Gram-Schmidt process.) (Strang, 1980). In practice, however, the QR decomposition is not performed by the Gram-Schmidt process because the Householder transformation technique allows the decomposition to be performed numerically with greater simplicity and stability (Strang, 1980).

The Householder transformation technique is a mapping that converts $\mathbf{A}$ into the sparse matrix $\mathbf{R}$, by using an orthonormal matrix $\mathbf{Q}^{-1} = \mathbf{Q}^T$ which satisfies the requirement that $\mathbf{Q}^{-1}\mathbf{A} = \mathbf{Q}^T\mathbf{A} = \mathbf{R}$. The matrix $\mathbf{Q}^{-1}$ is the product of the M orthonormal matrices $\mathbf{Q}_1, \mathbf{Q}_2, \ldots, \mathbf{Q}_M$ which individually define a Householder transformation (i.e. $\mathbf{Q}^{-1} = \mathbf{Q}_M, \ldots, \mathbf{Q}_2, \mathbf{Q}_1$). To understand how $\mathbf{Q}^{-1}$ maps $\mathbf{A}$ into the spare matrix $\mathbf{R}$, first note that the general Householder transformation is a vector operation that operates on a vector $\mathbf{x}$ to produce another vector which is the vector difference between the projection of $\mathbf{x}$ onto a arbitrary vector $\mathbf{b}$, denoted $P_{\mathbf{b}}(\mathbf{x})$, and the projection, $P_{\mathbf{b}}^{\perp}(\mathbf{x})$, of $\mathbf{x}$ onto a vector orthogonal to $\mathbf{b}$. Thus the Householder transformation of a vector $\mathbf{x}$ is given by the expression

$$Q_{\mathbf{b}}(\mathbf{x}) = P_{\mathbf{b}}^{\perp}(\mathbf{x}) - P_{\mathbf{b}}(\mathbf{x}), \tag{4.38}$$

which should be compared to the expression

$$\mathbf{x} = P_{\mathbf{b}}(\mathbf{x}) + P_{\mathbf{b}}^{\perp}(\mathbf{x}),$$

that describes $\mathbf{x}$ in terms of an orthogonal decomposition (Steinhardt, 1988)(see Fig. 4.19). In Fig. 4.19 the Householder transformation of $\mathbf{x}$ and an arbitrary vector $\mathbf{b}$ is illustrated pictorially

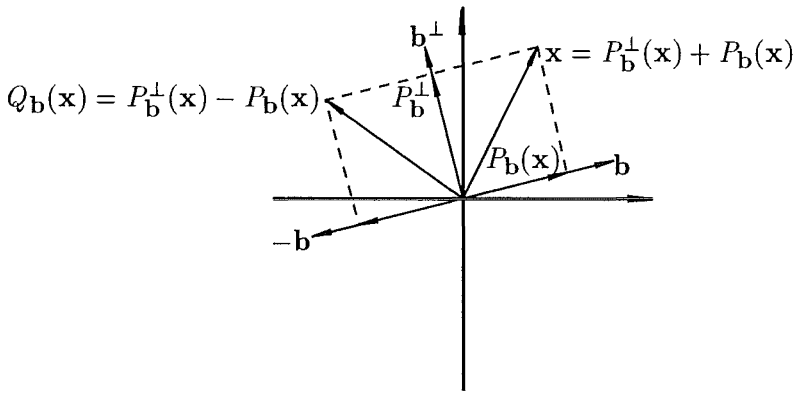

Figure 4.19 $Q_{\mathbf{b}}(\mathbf{x})$ the Householder transformation of $\mathbf{x}$ and the vectors which form the orthogonal decomposition of $\mathbf{x}$.

and it can be seen that $Q_{\mathbf{b}}(\mathbf{x})$ is simply the reflection of $\mathbf{x}$ in $\mathbf{b}^{\perp}$. This property is particularly important as it implies that the Householder transformation is an orthonormal operation since it preserves the length of the vectors on which it operates (Steinhardt, 1988).

The Householder or reflection matrix $Q_{\mathbf{b}}$, which when multiplied by a vector produces the Householder transformation of that vector, can be derived from Eqn.(4.38). To obtain $Q_{\mathbf{b}}$ the expressions $\mathbf{b} \cdot \mathbf{x} = |\mathbf{b}||\mathbf{x}| \cos\theta$, where θ is the angle between the two vectors (see Fig. 4.19) and $P_{\mathbf{b}}^{\perp}(\mathbf{x}) = \mathbf{x} - P_{\mathbf{b}}(\mathbf{x})$ are used to express $P_{\mathbf{b}}(\mathbf{x})$ and $P_{\mathbf{b}}^{\perp}(\mathbf{x})$ in the form

$$P_{\mathbf{b}}(\mathbf{x}) = \frac{\mathbf{b}\mathbf{b}^T}{\mathbf{b}^T\mathbf{b}}\mathbf{x}$$

and

$$P_{\mathbf{b}}^{\perp}(\mathbf{x}) = \mathbf{x} - \frac{\mathbf{b}\mathbf{b}^T}{\mathbf{b}^T\mathbf{b}}\mathbf{x}.$$

On substituting the above to expressions into Eqn.(4.38) the expression

$$Q_{\mathbf{b}}(\mathbf{x}) = \left(\mathbf{I} - 2\frac{\mathbf{b}\mathbf{b}^T}{\mathbf{b}^T\mathbf{b}}\right)\mathbf{x}$$

is obtained, and it can therefore be seen that $Q_{\mathbf{b}}$ is defined by (Rader and Steinhardt, 1986; Rader and Steinhardt, 1988)

$$Q_{\mathbf{b}} = \left(\mathbf{I} - 2\frac{\mathbf{b}\mathbf{b}^T}{\mathbf{b}^T\mathbf{b}}\right). \tag{4.39}$$

As the Householder transformation is a length-preserving or orthonormal process, the matrix $Q_{\mathbf{b}}$ must also be orthonormal. One of the most useful forms of Householder transformation is that where $\mathbf{b}$ is defined by the expression

$$\mathbf{b} = \mathbf{x} + |\mathbf{x}|\mathbf{e}_k, \tag{4.40}$$

where $\mathbf{e}_k$ is a standard basis vector in the K-D vector space spanned by $\mathbf{x}$ (e.g. $\mathbf{e}_1 = [1, 0, \ldots, 0]$ and $\mathbf{e}_K = [0, 0, \ldots, 1]$). In this situation it can be seen on substituting Eqn.(4.40) into Eqn.(4.39) and multiplying by $\mathbf{x}$ that

$$Q_{\mathbf{b}}\mathbf{x} = \mathbf{x} - 2\frac{(\mathbf{x} + |\mathbf{x}|\mathbf{e}_k)(\mathbf{x}^T + |\mathbf{x}|\mathbf{e}_k^T)}{(\mathbf{x}^T + |\mathbf{x}|\mathbf{e}_k^T)(\mathbf{x} + |\mathbf{x}|\mathbf{e}_k)}\mathbf{x}$$

$$
\begin{aligned}
&= \mathbf{x} - 2\frac{(\mathbf{x} + |\mathbf{x}|\mathbf{e}_k)(|\mathbf{x}|^2 + |\mathbf{x}|\mathbf{e}_k^T\mathbf{x})}{2|\mathbf{x}|^2 + 2|\mathbf{x}|\mathbf{e}_k^T\mathbf{x}} \\
&= -|\mathbf{x}|\mathbf{e}_k,
\end{aligned}
$$

consequently $\mathbf{x}$ has been reduced to the vector $-|\mathbf{x}|\mathbf{e}_k$ whose elements, except for the kth element, are all zero. In Steinhardt (1988) it is shown graphically that when $\mathbf{b}$ has the form described in Eqn.(4.40), the Householder transformation simply reflects the vector being transformed onto the standard basis vector that is a component of $\mathbf{b}$.

The reason for the utility of the above form of Householder matrix, is that when this transform is repeatedly applied to the result of the transform operating on a matrix, it is possible (if the $\mathbf{b}$ are chosen appropriately) to reduce the matrix to a sparse matrix. This is precisely the process which is required to reduce the coefficient matrix $\mathbf{A}$ in Eqn.(4.33) into the upper triangular matrix $\mathbf{R}$ and for this reason the application of the Householder transform to the QR decomposition problem can now be described (Dongarra *et al.*, 1979; Strang, 1980; Press *et al.*, 1986). On writing $\mathbf{A}$ in the form $[\mathbf{a}_{c1}, \mathbf{a}_{c2}, \ldots, \mathbf{a}_{cP}]$ and letting $\mathbf{b}_1 = \mathbf{a}_{c1} + |\mathbf{a}_{c1}|\mathbf{e}_1$ it can be seen that the effect of the Householder transform $\mathbf{Q}_1$ on $\mathbf{A}$ is to replace the entire first column of the matrix, except for the first element, with zeros. In order to perform the QR decomposition another matrix $\mathbf{Q}_2$ must now be applied to $\mathbf{Q}_1\mathbf{A}$ in order to replace all the elements below the second element in the second column with zeros. To accomplish this it is necessary to form another Householder matrix $\mathbf{Q}_2$ that when multiplied with $\mathbf{Q}_1\mathbf{A}$ reduces the second column of $\mathbf{A}$ without undoing the reduction that was previously performed by $\mathbf{Q}_1$. This second matrix $\mathbf{Q}_2$ is formed by inserting the $M-1 \times M-1$ Householder matrix $\tilde{\mathbf{Q}}_2$, obtained from the $M-1$ components of $\mathbf{Q}_1\mathbf{a}_{c2}$ below its first component, into the lower right hand corner of a $M \times M$ matrix whose first row and column elements, except for a one in the leading diagonal element, are all zero (Dongarra *et al.*, 1979; Strang, 1980; Press *et al.*, 1986), i.e.

$$
\mathbf{Q}_2 = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \tilde{\mathbf{Q}}_2 & \\ 0 & & & \end{bmatrix}.
$$

It can be seen that $\mathbf{Q}_2$, formed in the above manner, does have the properties required to reduce $\mathbf{Q}_1\mathbf{A}$ in the manner required. It is reasonably easy to see that by repeatedly applying the above process it is possible to reduce $\mathbf{A}$ into the upper triangular matrix $\mathbf{R}$. The matrix $\mathbf{Q}$ which is the other component of the QR decomposition can also be found quite simply by multiplying together all the Householder matrices, since $[\mathbf{Q}_M \cdots \mathbf{Q}_2\mathbf{Q}_1]\mathbf{A} = \mathbf{R}$ and therefore $\mathbf{Q}^{-1} = [\mathbf{Q}_M \cdots \mathbf{Q}_2\mathbf{Q}_1]$. The above development demonstrates that the Householder transform, which is numerically stable because it is simply a reflection, can indeed be used to decompose a linearly independent matrix $\mathbf{A}$ into the two matrices $\mathbf{Q}$ and $\mathbf{R}$.

4.5.3 Comparison of Reconstruction Techniques

In this section the Fourier-based and algebraic techniques of reconstructing an image are compared. This comparison aims to give the reader some feel for the flexibility, computational cost, and accuracy of the two techniques.

A major advantage that the algebraic technique has over the Fourier technique is that the zero samples required to obtain the reconstruction can be taken from any part of the image's zero sheet. The reason for this is that, when an image has been blurred and/or contaminated,

a superior reconstruction is most likely to be obtained if specific samples can be selected from the image's zero sheet because, as was shown in §4.4, contamination can perturb different regions of the zero sheet differently and in the deblurring case there may be some uncertainty as to whether a particular zero belongs to the psf or image (e.g. in the regions where merging has occurred). With the Fourier technique the zero samples selected must correspond to those zeros which allow the Fourier spectrum to be reconstructed along the lines of constant u or v in the real (u, v) plane (see §4.5.1). Consequently, if some of the zeros required by the Fourier technique happen to lie in the regions on the zero sheet that have been significantly perturbed by the contamination or in regions where merging has occurred, the reconstruction may not be the best possible reconstruction that could be obtained from the zero sheet. It is worth noting at this point that Satherley and Parker (1993) have recently developed an improved version of the Fourier technique which reconstructs the image's spectrum on complex planes in 4-D Fourier space (i.e. on other than the real (u, v) plane) from which an estimate of the image can be obtained after some further processing. Whilst this new extended technique provides some added flexibility, it still requires samples which lead to regular arrays of Fourier coefficients and thus lacks the flexibility of the algebraic technique. In summary, the ability to select zeros from anywhere on the zero sheet is a distinct advantage when an image is to be reconstructed from a contaminated and/or the blurred version of its zero sheet. The greater flexibility of the algebraic technique, compared to the Fourier technique, offers an advantage in these situations.

The flexibility of the algebraic technique over the Fourier technique comes at a cost; the computational requirements of the former technique exceed those of the latter technique quite significantly. In fact, it is shown that as the image size becomes large, the computational effort required to obtain a reconstruction may prohibit its use. In order to obtain a reconstruction by the algebraic technique the system of equations must first be formed; to form this system of M equations in P unknowns requires approximately

$$MP\sqrt{P}, \tag{4.41}$$

multiplications. To decompose the matrix $\mathbf{A}$ representing the system of equations into QR form, via Householder transformations, requires

$$MP - \frac{P^3}{3}, \tag{4.42}$$

multiplications and additions, and to obtain a solution to the system of equations by back substitution requires a further

$$2(M - P)P + \frac{1}{2}P^2 \tag{4.43}$$

multiplications and additions (Dongarra *et al.*, 1979). In the situation where a $N \times N$ image is to be uniquely reconstructed from its zero sheet $M = 2(N - 1)^2 + 1$ and $P = N^2 - 1$. Substituting the above values for M and P into Eqn.(4.41), Eqn.(4.42), and Eqn.(4.43) it can be seen that as N becomes large the computational cost of the algebraic technique scales $O(N^6)$ with the QR decomposition section of the reconstruction process being computationally most expensive.

With the Fourier technique the system of 1-D polynomials describing the spectrum of the image must first be derived and the inverse fast Fourier transform (Brigham, 1988) used to obtain an estimate of the image. As described in §4.5.1, $N+1$ polynomials that characterise the image's spectrum along the N lines $u = u_n$ and along one line $v = v_m$, must be evaluated N times and each evaluation requires $N - 1$ multiplications. Therefore to obtain the Fourier coefficients required by the IFFT requires $N(N - 1)(N + 1) \approx N^3$ multiplications. The IFFT must

perform $N^2 \log_2 N$ further multiplications to obtain an estimate of the image (Brigham, 1988). Thus, as the image size increases the computational cost of the Fourier technique is $O(N^3)$. It can therefore be seen from this comparison that the Fourier technique is significantly faster than the algebraic technique (i.e. $O(N^3)$ *vs* $O(N^6)$ multiplications). For large images it is very likely that the computational cost will preclude the use of the algebraic technique, hence, in this situation the much faster Fourier technique would have to be used.

The final attribute of the two reconstruction algorithms to be compared here is the reconstruction accuracy. The reconstruction accuracy of both techniques is compared here, somewhat subjectively, under the conditions where the set of zero samples used to obtain a reconstruction are: uncontaminated, contaminated, and erroneous (to the extent that one zero sample belongs to another image, i.e. one belongs to the psf in attempting to deconvolve). First consider the three cases for the Fourier technique.

In the situation where the IFFT is used to reconstruct a $N \times N$ image from a set of uncontaminated zero samples the error on the reconstruction is primarily due to the rounding errors that occur during the process of computing the IFFT. When a floating point implementation is used to reconstruct an image from its zero sheet the relative error due to this rounding is given by an expression of the form

$$\mathrm{SNR}_f = \frac{1}{N \log_2 N \sigma^2 \epsilon},$$

where SNR_f is the SNR of the reconstructed image and $\sigma^2\epsilon$ is the variance of the white noise source used to model the rounding error in the computer used (Oppenheim and Schafer, 1975). It can be seen from the above expression that the Fourier technique should be able to reconstruct an image to an extremely good accuracy since $\sigma^2\epsilon$ for most modern computers, which have four byte and/or eight byte real number representation schemes, should be very small. As the size of the image increases it can also be seen that the error in the reconstruction due to rounding error increases, but does so at a reasonably slow rate.

Now consider the case where the zero sheet has been corrupted by contamination of a level sufficient to ensure that the majority of the error on the reconstruction is due to the zero sheet contamination, but not so severe that the set of zeros required to obtain a reconstruction, via the Fourier technique, cannot be correctly identified. In this situation the set of zeros used to reconstruct the spectrum can be thought of as slightly perturbed versions of the uncontaminated zeros. The level of contamination on the reconstructed spectrum is therefore going to be proportional to the size of the perturbations on these zeros and the level of contamination on the final image is going to be equal to that on the reconstructed spectrum. Thus, in the above situation the error on the reconstructed image is going to be proportional to the level of contamination on the zero sheet and this error should not be directly related to image size.

When one of the zero samples used to reconstruct the Fourier spectrum belongs to another image, which could occur when zero sheet separation techniques are being used to deblur an image, the accuracy of the reconstruction can be very poor. This is because if one zero sample is completely incorrect then the level of contamination on the reconstructed spectrum, due to this rogue zero sample, is completely dependent on how close this rogue sample is to the correct sample. There is no guarantee that the reconstructed spectrum has the correct form (e.g. if the image is a real image it reconstructed spectrum may no longer be conjugate symmetric) and the level of contamination on the spectrum may be very large. Thus, in this case it is very likely that the Fourier technique would be unable to obtain a useful reconstruction. In §5.5 some experimental results are presented which demonstrate that Fourier reconstruction is indeed very sensitive to the incorrect selection of zero samples.

When QR decomposition and Householder transforms are used to obtain an algebraic reconstruction of an image from its zero sheet, in the uncontaminated situation, the rounding errors in the computer used result in the algorithm minimising an equation of the form

$$|(\mathbf{A}+\mathbf{F})\mathbf{x}-\mathbf{y}|^2,$$

which should be compared to Eqn.(4.32) (Dongarra *et al.*, 1979). In this situation the size of the perturbations on $\mathbf{A}$, which are represented here by $\mathbf{F}$, are given by an expression of the form

$$\frac{|\mathbf{F}|_F}{|\mathbf{A}|_F} \leq \Upsilon(M,P)\epsilon_m,$$

where $|\mathbf{A}|_F^2 = \sum_{i,j} A_{i,j}^2$ is the Frobenius matrix norm, $\Upsilon(M,P)$ is a small slowly varying function of M and P that grows significantly slower than M^2P^2, and ϵ_m is the rounding unit of the computer used (Dongarra *et al.*, 1979). It can therefore be seen, as was the case with the Fourier technique, that the algebraic technique should be able to reconstruct images to a reasonably good accuracy in the uncontaminated case. As the image size increases the increase in the error is significantly less than $O(N^8)$ (i.e. $M^2P^2 = ((2(N-1)^2+1)^2(N^2-1)^2)$, which could be much greater than that of the Fourier technique. In §5.5 a set of experimental results demonstrate that in the uncontaminated case the Fourier and algebraic techniques can provide reconstructions of comparable accuracy.

In the situation where the image's zero sheet is contaminated, the coefficient matrix $\mathbf{A}$ is perturbed in a manner that is proportional to the level of contamination on the zero sheet. When the coefficient matrix has the form $\mathbf{A}+\mathbf{E}$, where $\mathbf{E}$ is a matrix representing the perturbations, the error in the least squares solution can be described by an equation of the form

$$\frac{|\mathbf{x}-\hat{\mathbf{x}}|}{\mathbf{x}} \leq 1.6\left[\kappa(\mathbf{A})+\kappa^2(\mathbf{A})\frac{|\mathbf{R}|}{|\mathbf{A}||\mathbf{x}|}\right]\frac{|\mathbf{E}|}{|\mathbf{A}|}, \tag{4.44}$$

where $\kappa(\mathbf{A}) = |\mathbf{A}||\mathbf{A}^{-1}| \approx \mathbf{R}_{00}/\mathbf{R}_{PP}$ is the condition number of the matrix $\mathbf{A}$ (Dongarra *et al.*, 1979). When a matrix is well conditioned the condition number is close to one and as the matrix becomes ill-conditioned the condition number increases. The expression in Eqn.(4.44) demonstrates that as the level of contamination increases on the zero sheet, the corresponding increase in $|\mathbf{E}|$ is going to result in a decrease in the accuracy of the solution. In the situation where the image gets larger the $2(N-1)^2$ zero samples that must be selected from the zero sheet get closer together in Z space (see §5.4) and consequently the linear independence of the set of equations comprising $\mathbf{A}$ decreases. When the linear independence of $\mathbf{A}$ decreases $\mathbf{R}_{PP}$ also decreases and therefore there is a corresponding increase in $\kappa(\mathbf{A})$ which results in a decrease in the reconstruction accuracy. Thus, as the image size increases there is a corresponding decrease in the accuracy of the reconstructions obtained by this technique. It is very likely that for large images the reconstruction accuracy of the Fourier technique could be better than that of the algebraic technique due to the ill-conditioning problem the latter technique is likely to have for larger images. For smaller images, however, if the contamination on the equations comprising $\mathbf{A}$ is Gaussian distributed it is likely that the least squares solution provided by the algebraic technique could be superior to that obtained from the Fourier technique.

Finally consider the situation in which a zero sample is incorrect. In this situation it is very likely that the algebraic technique could still obtain a useful reconstruction. The reason for this is that the residual associated with the equation derived from this rogue zero sample is only one term in the summation which is minimised to obtain a least squares solution and a reconstruction. As the least squares solution aims to minimise the total of all the residuals,

the residuals associated with the $2(N-1)^2$ correct zero samples are going to be very small. Consequently, the solution obtained by the algebraic technique, whilst it may not be as accurate as that which would be obtained if the equation associated with the rogue sample was removed, is likely to be of a sufficiently good accuracy to be useful. In summary, the reconstruction accuracy of both reconstruction techniques is likely to be very similar in the uncontaminated situation. In the contaminated situation, where all the zero samples are from the correct zero sheet, the least squares solution provided by the algebraic technique is likely to be superior to that from the Fourier technique, however as the image size increases this situation is likely to reverse. When one of the zero samples is incorrect the algebraic technique is the most likely of the two techniques to provide a useful reconstruction.

CHAPTER 5

2-D ZERO-BASED CONSISTENT DECONVOLUTION

In §3.2 it was shown that the accuracy of image estimates provided by the existing 2-D standard deconvolution techniques is heavily dependent on the accuracy of the psf estimates. It was also shown in §3.2.4 and §3.4 that by using the consistent deconvolution principle, it should be possible to develop a standard deconvolution algorithm which is considerably less sensitive to the psf accuracy than the existing 2-D techniques. In this chapter a new algorithm that uses the consistent deconvolution principle to solve the contaminated standard deconvolution problem is described. The new algorithm is known as the 2-D consistent deconvolution algorithm. The 2-D consistent deconvolution algorithm is an extension of the 1-D consistent deconvolution algorithm of Bates *et al.* (1990) (see §3.2.4) and the algorithms of Lane *et al.* (1987) and Lane and Bates (1987) which use zero sheets to solve either the uncontaminated blind deconvolution or phase retrieval problem. The novel features of this new algorithm are: (i) it uses zero sheets to solve the 2-D standard deconvolution problem, (ii) can solve the 2-D standard deconvolution problem in a consistent manner, and (iii) is the first algorithm to successfully use zero sheets to solve the contaminated 2-D standard deconvolution problem. It is drawn to the reader's attention that the zero sheets of 1-D images (the 'sheets' degenerate into point zeros) are distinctly different to those of multidimensional images (see §4.1) and that the zero sheets of uncontaminated blurred images undergo major changes when the images are contaminated (see §4.4). The motivation for developing the 2-D consistent deconvolution algorithm was that it was anticipated that the algorithm's accuracy should be less sensitive to the psf accuracy than the existing 2-D techniques described in §3.2 (like the 1-D consistent deconvolution technique, the 2-D consistent deconvolution technique only uses the psf for recognition purposes and does not use $\hat{h}(x, y)$ explicitly to compute $\hat{f}(x, y)$).

The first section of this chapter contains an overview of the 2-D consistent deconvolution algorithm. In this first section the algorithm is divided into its major functional elements and the principle behind each of these components is discussed. In §5.2 and §5.3 the implementation of the two most important functional elements within the algorithm are described in detail. The first of these elements is used to perform the recognition process that allows the zero sheet of the image to be recognised in, and separated from, the zero sheet of the blurred image. Experimental results that demonstrate the effectiveness of the above process are also presented in §5.2. In §5.3 the part of the algorithm responsible for selecting the samples from the image estimate's zero sheet and obtaining a reconstruction from these samples is described. This section primarily focuses on the automatic zero sheet sampling techniques that should be used to ensure that good estimates of the original image are constantly obtained when the blurred image has been corrupted with wide-band or band-limited contamination. Results from simulations are presented in §5.3 to demonstrate the capabilities of the sampling techniques described.

In §5.4 the effects of psf size and image size on the accuracy of the $\hat{f}(x,y)$ provided by the 2-D consistent deconvolution are described with supporting experimental results. The dependence of $\hat{f}(x,y)$ Q on the SNR of the blurred image is also described in §5.4. The 2-D consistent deconvolution algorithm reconstructs the $\hat{f}(x,y)$ from its zero sheet by using the algebraic reconstruction technique (see §4.5.2). In §5.5 the results obtained by using the Fourier reconstruction technique are compared to those obtained by using algebraic reconstruction. The computational requirements of the 2-D consistent deconvolution algorithm are described in §5.6 in terms of the amount of CPU time required to obtain an $\hat{f}(x,y)$. The effect that image size has on the computational effort required is also described in this section.

The performance of the 2-D consistent deconvolution algorithm has been compared to that of the Wiener filter and the results of this comparison are presented in §5.7. This comparison covers images of various sizes, blurred images of various SNRs, and $\hat{h}(x,y)$ of various accuracies. The new algorithm is compared to the Wiener filter as both algorithms are direct frequency domain techniques and the performance of the Wiener filter is well understood. Recently another 2-D consistent deconvolution algorithm, which uses a different zero-based matching technique, has been developed by Parker (1992). As with the Wiener filter the performance of the algorithm of Parker (1992) has been compared to that of the 2-D consistent algorithm described here and the results from this comparison are presented in §5.8. Finally, in §5.9 the author describes the difficulties which are anticipated if the algorithm was used to deconvolve practical images. Possible techniques which could be used to overcome some of these difficulties are also described.

5.1 OVERVIEW OF ALGORITHM

In this section the main object of the new 2-D consistent deconvolution algorithm is described. In addition, the purpose of the major components within the algorithm are overviewed and the principles that provide the basis for each of these components are discussed.

It was shown in §3.2 that all the existing 2-D standard deconvolution techniques suffer from the limitation that the accuracy of the $\hat{f}(x,y)$ that they provide is strongly related to the accuracy of $\hat{h}(x,y)$. This limitation was the main motivation for developing the new 2-D consistent deconvolution algorithm which is described here. It was the author's expectation that the new 2-D consistent deconvolution algorithm, known hereafter as the consistent algorithm, could use the consistent deconvolution principle (see §3.4) to obtain $\hat{f}(x,y)$ that have an accuracy that is not explicitly related to the psf accuracy. The consistent deconvolution principle uses $\hat{h}(x,y)$ to recognise the actual psf responsible for the blurring and this in turn allows the actual blurring to be removed from the blurred image. Thus, the consistent algorithm only uses $\hat{h}(x,y)$ for recognition purposes and does not explicitly use $\hat{h}(x,y)$ to compute $\hat{f}(x,y)$. For this reason the consistent algorithm should be less susceptible to psf inaccuracies than the existing techniques.

The consistent algorithm implements the consistency condition by using the property that a blurred and contaminated image's zero sheet, for modest levels of contamination, is approximately formed from the union of the zero sheets pertaining to the psf and original image (see §4.4.3). For moderate levels of contamination on the convolution (i.e. $\leq$ 30dB, see §4.4.3) it is therefore possible to use $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$ to recognise $Z\{\mathcal{H}(\zeta,\gamma)\}$ within $Z\{\mathcal{G}(\zeta,\gamma)\}$. On removing $Z\{\mathcal{H}(\zeta,\gamma)\}$ from $Z\{\mathcal{G}(\zeta,\gamma)\}$ the best possible estimate of $Z\{\mathcal{F}(\zeta,\gamma)\}$ remains and it is possible to reconstruct an estimate of this image from the remaining zero sheet. In the uncontaminated situation, a consistent solution to the standard deconvolution problem can be

obtained since the union relationship between the zero sheets is exact.

The main sections of the consistent algorithm are now outlined. To obtain a solution to the standard deconvolution problem the consistent algorithm sequentially implements the following operations:

- The required zero sheet slice, or slices, pertaining to $g(x,y)$ and $\hat{h}(x,y)$ are computed.
- Those sections of $Z\{\mathcal{G}(\zeta,\gamma)\}$ that match those of $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$ are determined, i.e the sections corresponding to $Z\{\mathcal{H}(\zeta,\gamma)\}$ are recognised.
- Those sections of $Z\{\mathcal{G}(\zeta,\gamma)\}$ corresponding to $Z\{\mathcal{H}(\zeta,\gamma)\}$ are removed and the remaining sections of $Z\{\mathcal{G}(\zeta,\gamma)\}$ are pieced (spliced) together to form $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$
- Point zeros are chosen from $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ in a manner that allows the best possible estimate of the original image to be obtained by the algebraic reconstruction technique.
- The image is reconstructed, to within the arbitrary complex constant, using the algebraic reconstruction technique.
- For the purpose of obtaining comparative results, the value of the arbitrary complex constant is determined and thereby the image estimate's quality factor.

Of the above operations those that required significant amounts of research and development were the recognition and separation, and sampling sections. The original research that describes the author's attempts to perfect these sections of the algorithm is described in §5.2 and §5.3 respectively. The reconstruction of the image estimate from its zero sheet samples by algebraic reconstruction is not described in this chapter as the algebraic reconstruction technique, and the QR decomposition algorithm used to solve the overdetermined system of equations, has been extensively described in §4.5.2. The QR decomposition routine used here was taken from the linear equations solution package LINPACK (Dongarra *et al.*, 1979) and adapted by the author to operate on complex double precision data (2×8 bytes) as opposed to real single precision data.

In the consistent algorithm the zero sheet slice or slices corresponding to the circular contour defined by Eqn.(4.6) are produced for preselected ρ values. In §5.3 it is shown how the value of ρ can be chosen to ensure that the consistent algorithm can reliably provide an $\hat{f}(x,y)$ of the best possible Q. The phase term, ϕ, in Eqn.(4.6) is incremented in 1024 equal angular increments between $-\pi$ and π. The zero sheet slice is then formed from the sets of zeros belonging to the resulting 1-D polynomials in the manner described in §4.2.1. The number of equal angular increments required between $-\pi$ and π is data dependent and 1024 was chosen for historic reasons as is outlined in §5.2. Note that 1024 angular increments in ϕ appears to be sufficient to guarantee that the zeros can be unambiguously linked (for images of small size, i.e. $\leq 32 \times 32$ pixels, as processed here) as described in §4.2.1. If the number of increments is made too small the zero sheets of a convolution would appear to merge in a manner that is similar to the genuine merging which occurs when a convolution is contaminated. As the merging due to convolution contamination and selecting an insufficient number of increments in ϕ is difficult to distinguish, it is anticipated that a successful zero sheet recognition and separation algorithm should be able to handle both forms of merging equally well.

The consistent algorithm uses the algorithm CPOLY (Jenkins and Traub, 1972) to compute the zeros of the 1-D polynomials that are produced as ζ is moved around the circular contour.

The CPOLY algorithm is a three stage algorithm that computes the zeros of a complex polynomial one zero at a time (Jenkins and Traub, 1970; Ralston and Rabinowitz, 1978). After finding a particular zero, the polynomial is deflated and the algorithm finds the next zero. The deflation is performed by removing the terms containing the zeros of smallest magnitude first as stability problems arise if the polynomial is deflated using the terms associated with the larger magnitude zeros. There would be little benefit to the reader in giving a detailed description of the three stages of the algorithm and so only a brief overview is given. Each of the three stages use a technique which is similar to the Newton-Raphson technique to find the smallest zero of the current polynomial. The technique converges to the zero of the polynomial which is closest to some predefined origin and in each stage the origin is adjusted in a different manner to allow the technique as a whole to converge on the zero of smallest magnitude. In the first stage of the algorithm the origin is set to zero and the algorithm begins to converge on a zero of small magnitude, the second stage then shifts the origin to an arbitrary region of Z space with a magnitude less than that of the smallest zero and this allows the algorithm to converge on a single dominant zero within any clusters of zeros, and in the final stage the origin is set to the latest estimate of the dominant zero (c.f. Newton-Raphson) which allows the algorithm to refine its estimate of this zero (Jenkins and Traub, 1970; Ralston and Rabinowitz, 1978). It is noteworthy that the final stage of the algorithm converges quadratically. Jenkins and Traub (1972) state that CPOLY is an extremely fast algorithm and that the computational effort required is not heavily dependent on the distribution of the zeros.

The image estimate obtained from the consistent algorithm differs from $f(x,y)$ by the inevitable contamination and the arbitrary complex constant. In practice it is not necessary to know the complex constant as the reconstructed image is simply an image form of $f(x,y)$, i.e. it has the same structure. When conducting simulations to establish how the algorithm performs in different situations and configurations, however, it is necessary to be able to compute this complex constant and the dependent quality factor (see §2.1). Thus, a technique has been developed to remove the complex constant from the reconstructed image estimate $\tilde{f}(x,y)$, using prior knowledge not usually available, to give a image estimate $\hat{f}(x,y)$ that only differs from $f(x,y)$ by residual error and contamination. This technique is now described and in the discussion that follows it is assumed that the original image is a real image. Complex images can be handled by the technique and the method used for complex images is described after that for real images.

To compute the complex constant $C\exp^{i\psi}$ (C and ψ real) by which $\tilde{f}(x,y)$ must be scaled to give $\hat{f}(x,y)$, prior knowledge of $\|f(x,y)\|^2$ is used first to determine C. The phase on the pixels of $\tilde{f}(x,y)$, $\chi(x,y)$, consists of the constant component ψ and an error $(\psi - \chi(x,y))$, which is due to the contamination on the convolution and the inaccuracies introduced by the reconstruction process. As the $(\psi - \chi(x,y))$ should be approximately randomly distributed with zero mean, the average phase of the pixel values should be a good estimate of ψ. It is also reasonable to expect that the $(\psi - \chi(x,y))$ associated with the pixels of large magnitude should be smaller than that of the pixels of small magnitude. A weighted average of the phases of the pixels, where the phases of the more accurate large magnitude pixels are weighted heavier than those of the small magnitude pixels, can therefore be used to determine ψ. This weighted average is computed by calculating the value of ψ which maximises the expression

$$\wp = \sum_{n=0}^{N-1} |\check{f}[n,m]| \cos(\chi[n,m] - \psi),$$

$$= \cos(\psi) \sum_{n,m=0}^{N-1} |\check{f}[n,m]| \cos(\chi[n,m]) + \sin(\psi) \sum_{n,m=0}^{N-1} |\check{f}[n,m]| \sin(\chi[n,m]). \quad (5.1)$$

The above expression describes the summation of the projections of the vectors describing the pixel values of $\tilde{f}(x,y)$ onto the vector $\exp^{i\psi}$. On differentiating Eqn.(5.1) with respect to ψ and equating the resulting expression to zero, the expression

$$\psi = \arctan \left\{ \frac{\sum_{n,m=0}^{N-1} |\check{f}[n,m]| \sin(\chi[n,m])}{\sum_{n,m=0}^{N-1} |\check{f}[n,m]| \cos(\chi[n,m])} \right\},$$

describing the optimum value of ψ is obtained. To handle complex images, the phase of $f(x,y)$, which is known *a priori*, is subtracted from $\tilde{f}(x,y)$ to effectively make $\tilde{f}(x,y)$ approximate a real image. After computing and removing the phase $\exp^{i\psi}$ from all the pixels of $\tilde{f}(x,y)$ by the above technique it is then possible to add the known phase of $f(x,y)$ back onto the image to obtain $\hat{f}(x,y)$.

5.2 RECOGNISING AND SEPARATING THE COMPONENT ZERO SHEETS

One of the most crucial sections of the consistent algorithm is the section which is devoted to obtaining $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ from $Z\{\mathcal{G}(\zeta,\gamma)\}$ by using the known $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$ to recognise and eliminate $Z\{\mathcal{H}(\zeta,\gamma)\}$ from $Z\{\mathcal{G}(\zeta,\gamma)\}$. The task is complicated by the distortion and coalescence of the component zero sheets caused by contamination. The algorithm developed was therefore required to meet the following specifications:

- The algorithm must take the same slice from $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$ and $Z\{\mathcal{G}(\zeta,\gamma)\}$, and recognise and remove those sections of zero track forming the $Z\{\mathcal{G}(\zeta,\gamma)\}$ slice that most closely resemble the $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$ slice. The algorithm must perform this operation in the presence of contamination-induced zero sheet merging.
- Where zero sheets merge the algorithm must recognise this and separate the component zero tracks forming the $Z\{\mathcal{G}(\zeta,\gamma)\}$ slice at an appropriate place.
- If unable to decide with a reasonable degree of certainty whether a part of $Z\{\mathcal{G}(\zeta,\gamma)\}$ belongs to $\mathcal{F}(\zeta,\gamma)$ or $\mathcal{H}(\zeta,\gamma)$, the algorithm must include this section within $Z\{\mathcal{H}(\zeta,\gamma)\}$ (Note that the complete zero sheet of $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ or slice of $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ is not required for the purposes of obtaining a reconstruction).
- Subsequently the algorithm must splice together the remaining sections of $Z\{\mathcal{G}(\zeta,\gamma)\}$ to give the best possible estimate of $Z\{\mathcal{F}(\zeta,\gamma)\}$.
- As an overall requirement the algorithm should be simple, fast and (most importantly) robust.

Two methods were originally considered as candidates for recognising and separating the component zero tracks in a $Z\{\mathcal{G}(\zeta,\gamma)\}$ slice. The first of these techniques represents each zero track of $Z\{\mathcal{G}(\zeta,\gamma)\}$ and $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$ by a set of Fourier descriptors (Persoon and King Sun, 1977). To form a set of Fourier descriptors the complex zeros that comprise each of the zero tracks of the above images are used as the input to a Fourier transform algorithm. The Fourier descriptors of the individual zero tracks are then the output from

each Fourier transform. This technique led to the use of 1024 angular increments in the circular contour used to produce the zero sheet slices (see §5.1), since a 1024 sample FFT was used to compute the Fourier descriptors. Preliminary results showed that by comparing the Fourier descriptors belonging to each zero track in $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$ to all the Fourier descriptors describing the zero tracks in $Z\{\mathcal{G}(\zeta,\gamma)\}$ it would be possible, using a minimum square error technique, to automatically recognise those tracks in $Z\{\mathcal{G}(\zeta,\gamma)\}$ that most closely resemble those comprising $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$. This method was not pursued further, however, as there did not appear to be any way of adapting the technique to accurately recognise sections of zero track that had merged or to provide information on the location of the merger.

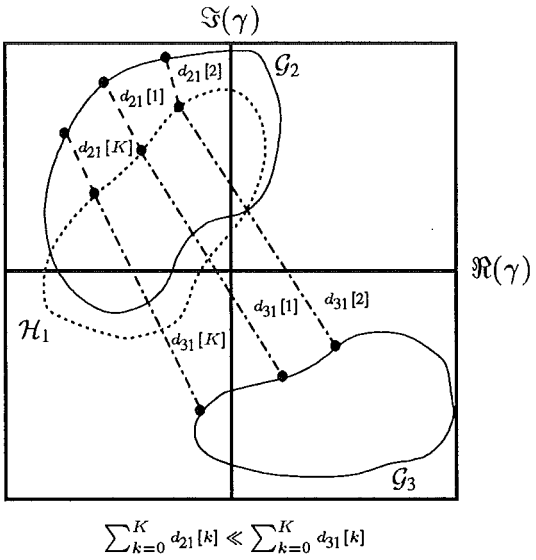

Figure 5.1 A schematic showing how the Euclidean error measure determines which zero tracks of the convolution belong to the psf. Shown is a psf track $\mathcal{H}_1$ which is being compared by the Euclidean error measure to two convolution tracks $\mathcal{G}_2$ and $\mathcal{G}_3$. As can be seen, the Euclidean error measure correctly identifies the convolution track $\mathcal{G}_2$ as belonging to the psf.

The second recognition technique, and that used to produce the results in Table 5.2 (No Splice) and Watson *et al.* (1992), uses a simple Euclidean error measure to determine which zero tracks of $Z\{\mathcal{G}(\zeta,\gamma)\}$, on a particular slice of the zero sheet, were closest in a Euclidean sense to those of $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$. The Euclidean error measure used was defined by the equation

$$\mathcal{E}_{pq} = \sum_{k=1}^{K} |Z_p\{\mathcal{G}(\zeta_k,\gamma)\} - Z_q\{\hat{\mathcal{H}}(\zeta_k,\gamma)\}| = \sum_{k=1}^{K} d_{pq}[k], \tag{5.2}$$

where the subscripts p and q denote that the qth zero track comprising the $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$ slice is being compared to the pth zero track comprising the $Z\{\mathcal{G}(\zeta,\gamma)\}$ slice (see Fig. 5.1). In Eqn.(5.2) the index k denotes the zero produced by setting the phase term, ϕ, in Eqn.(4.6) to $\frac{2\pi k}{K}$, where $k = 1,\ldots,K$ and here K = 1024. For each value of q the zero track of $g(x,y)$ that results in the minimum value of $\mathcal{E}_{pq}$ is the zero track that is considered to be the best estimate of the qth zero track of $\hat{h}(x,y)$ (see Fig. 5.1). Eliminating these q zero tracks from the $Z\{\mathcal{G}(\zeta,\gamma)\}$ slice should leave the $p-q$ zero tracks of $g(x,y)$ that give the best possible $\hat{f}(x,y)$. It is explained latter how this simple and robust technique can be adapted so that it can recognise and separate merged tracks. This basic recognition technique was incorporated into a version of the consistent algorithm, which operated on a single psf and convolution zero sheet slice and sampled the resulting $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ slice using the techniques in §5.3, to provide

an algorithm that could solve the standard deconvolution problem (Watson and Bones, 1991; Watson *et al.*, 1992).

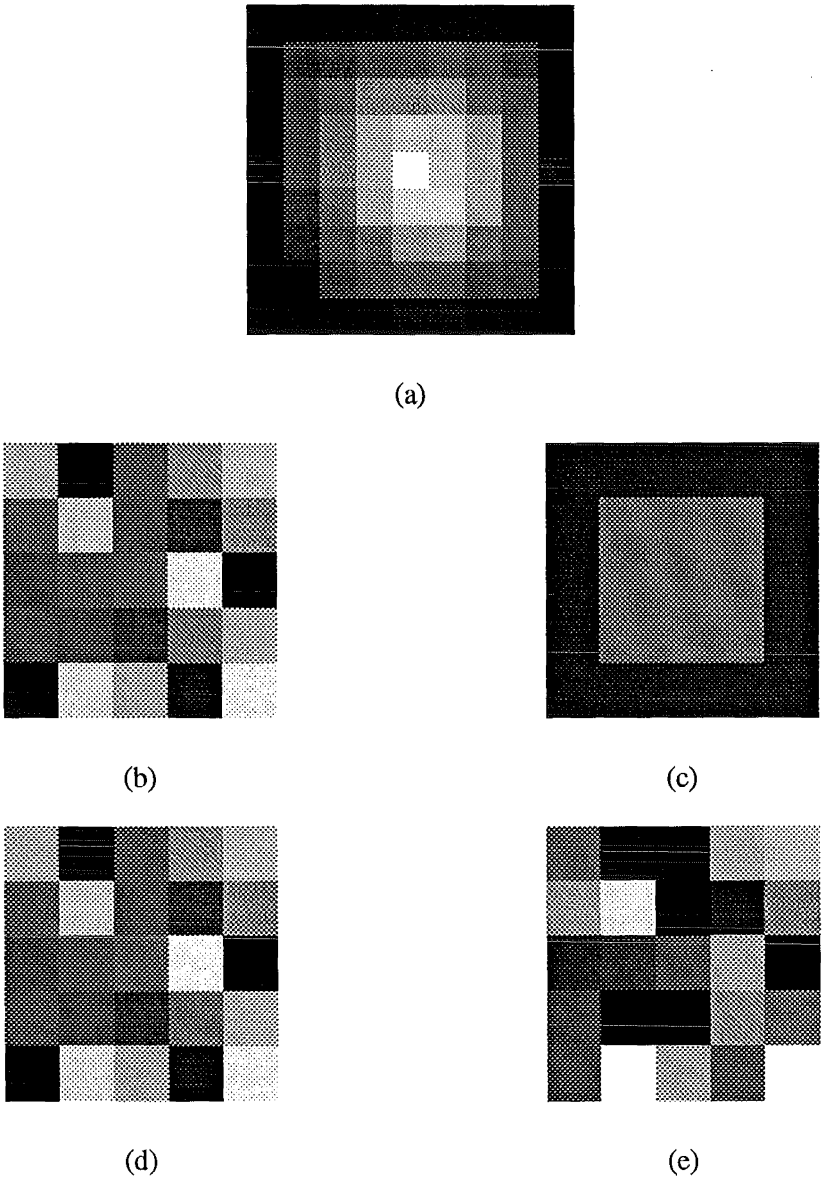

Figure 5.2 (a) The convolution of the random 5×5 image in (b) with the bilevel circular 5×5 psf in (c). The reconstructions obtained using the consistent algorithm on 60 dB and 40 dB SNR versions of (a) are shown in (d) (Q = 35.4 dB) and (e) (Q = 8.0 dB) respectively.

To confirm that this algorithm could provide a solution to the uncontaminated deconvolution problem and to determine whether the algorithm could obtain solutions in the situation where the convolution had a high SNR, a set of preliminary deconvolution results were obtained. The algorithm successfully deconvolved uncontaminated and contaminated versions of the blurred image in Fig. 5.2 (a), produced by convolving the 5×5 pixel random image in Fig. 5.2 (b) with the bilevel psf in Fig. 5.2 (c). Various slices (i.e values of the parameter ρ) were employed. In the situation where the blurred image was uncontaminated and $h(x, y)$ was used as the psf estimate, it was found that the algorithm provided $\hat{f}(x, y)$ with Qs of typically 130 dB. These results indicated that, in the uncontaminated situation, the algorithm was able to obtain

reconstructions of an accuracy commensurate with the numerical precision of the computer used.

To determine whether the algorithm could obtain good quality image estimates in the contaminated situation the blurred image in Fig. 5.2 (a) was contaminated with Gaussian distributed noise and $\hat{f}(x,y)$ were obtained from these convolutions. The contamination was added to the blurred image using an approach that gave the resulting convolution either a 60 dB or 40 dB SNR. The quality of the $\hat{f}(x,y)$ obtained from both the 60 dB and 40 dB SNR convolutions are tabulated in Table 5.2, under the column heading 'No splice', for slices corresponding to ρ values ranging from 0.6 to 1.1. For the 60 dB SNR convolution, the quality of the $\hat{f}(x,y)$ obtained ranged from 2.4 dB at $\rho = 1.0$ to 35.4 dB at $\rho = 1.1$ and for the 40 dB SNR convolution the $\hat{f}(x,y)$ Q ranged from 0.6 dB at $\rho = 1.0$ to 8.0 dB at $\rho = 0.7$. The best reconstructions obtained with this Euclidean-based algorithm from the 60 dB SNR and 40 dB SNR convolution are shown in Fig. 5.2 (d) and Fig. 5.2 (e) respectively. The sensitivity of the $\hat{f}(x,y)$ Q to the value of ρ is a consequence of the localised nature of the zero sheet merging (see §4.4.3). That is, if the selected slice does not correspond to a slice that incorporates a localised region of zero sheet merging, the component zero tracks have not coalesced in that slice and as a result the Euclidean based algorithm can obtain an accurate solution. As explained in §4.4.3 the zero sheet merging tends to decrease as $\rho \rightarrow \infty$ and $\rho \rightarrow 0$ because the component zero sheets tend to shrink towards a set of point zeros. For this reason the image estimates for $0.9 < \rho > 1.0$ are much better than those for $0.9 \geq \rho \leq 1.0$; in the latter the component zero sheets tend to have their maximum extent and tend to merge more often.

As shown above, the algorithm can obtain reconstructions of an accuracy commensurate with the numerical precision of the computer used. To determine whether the algorithm can still accomplish this when $\hat{h}(x,y)$ is inaccurate (i.e. obtain a consistent solution), random contamination was added to $h(x,y)$ in a manner which produced $\hat{h}(x,y)$ with SNRs of 60 dB, 50 dB, 40 dB, 30 dB and 20 dB. This series of psf estimates were then used by the algorithm to deconvolve a number of 12×12 and 20×20 convolutions. To demonstrate that the inverse filter is unable to obtain consistent solutions to the uncontaminated deconvolution problem, the images were also deconvolved by an inverse filter using the above set of $\hat{h}(x,y)$. The results presented in Table 5.1 demonstrate that the algorithm can provide consistent solutions to the standard deconvolution problem, as the $\hat{f}(x,y)$ Q are comparable to the machine's accuracy and only fall when $Z\{\mathcal{H}(\zeta,\gamma)\}$ can no longer be accurately recognised in $Z\{\mathcal{G}(\zeta,\gamma)\}$. In contrast, the quality of the $\hat{f}(x,y)$ obtained from the inverse filter are strongly dependent on the $\hat{h}(x,y)$ accuracy, indicating that this filter is unable to provide a consistent solution to the standard deconvolution problem.

RECONSTRUCTIONS CONFIRMING THE CONSISTENT SOLUTION

	20×20 $g(x,y)$		12×12 $g(x,y)$	
psf SNR dB	Q dB Zeros	Q dB Inverse	Q dB Zeros	Q dB Inverse
60	111.81	48.90	118.11	47.50
50	111.23	33.80	117.53	32.80
40	111.40	22.90	118.04	29.20
30	84.77	9.44	116.84	16.10
20	79.78	12.40	85.14	16.20

Table 5.1 Zeros = $\hat{f}(x,y)$ obtained from consistent algorithm. Inverse = $\hat{f}(x,y)$ obtained from inverse filter.

To satisfy the requirements of the specification, and to allow contaminated blurred images larger than 9×9 pixels to be deconvolved by the consistent algorithm, it was necessary to extend the Euclidean based algorithm so that it could recognise and separate sections of zero track that had merged together (Watson and Bones, 1992). The extension of the algorithm was based around two observations of the behaviour of the contaminated zero sheet. The first was the observation of Bates *et al.* (1990) that for reasonable levels of contamination the zero tracks of $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$ tend to head in the same direction in Z space as those zero tracks of $Z\{\mathcal{G}(\zeta,\gamma)\}$ which correspond to $Z\{\mathcal{H}(\zeta,\gamma)\}$. The second observation, noted by the author, was that if the distance between a zero track in $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$ and its corresponding zero track in $Z\{\mathcal{G}(\zeta,\gamma)\}$ is plotted as a function of ϕ, the distance between the two tracks dramatically increases when the zero track in $Z\{\mathcal{G}(\zeta,\gamma)\}$ has merged with another track. When this occurs there is a corresponding decrease in the distance between the psf track being followed and one of the other zero tracks in $Z\{\mathcal{G}(\zeta,\gamma)\}$, indicating that this other zero track in $Z\{\mathcal{G}(\zeta,\gamma)\}$ is the best estimate of the remainder of the zero track being followed. The manner in which the above principles have been incorporated within the basic Euclidean recognition scheme to give a technique that can separate merged zero tracks is now described.

As with the basic Euclidean recognition technique the improved technique uses the simple Euclidean distance measure (Eqn.(5.2)) to initially determine which of the zero tracks of $Z\{\mathcal{G}(\zeta,\gamma)\}$ best match those of $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$. To allow the zero tracks of $Z\{\mathcal{G}(\zeta,\gamma)\}$ to be spliced together the principles outlined above are incorporated into the algorithm by introducing two further measures which are computed for the K zeros that form each of the zero tracks of $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$. The first measure, denoted $d_{pq}[k]$, is the distance between the kth zero forming the qth zero track of $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$ and the kth zero forming the pth zero track of $Z\{\mathcal{G}(\zeta,\gamma)\}$. At each value of k a reference to the zero track of $Z\{\mathcal{G}(\zeta,\gamma)\}$ which is closest to the qth zero track of $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$ (i.e. that which resulted in the smallest $d_{pq}[k]$) is then stored in the first of two reference arrays for the qth zero track of $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$. As an example, the first reference array that results from the comparison of the psf estimate zero track $\hat{\mathcal{H}}_1$ to the convolution zero tracks $\mathcal{G}_1$, $\mathcal{G}_2$ and $\mathcal{G}_3$ in Fig. 5.3 is also shown in Fig. 5.3. According to the figure, the first section (k=1 to k=i+2) of $\hat{\mathcal{H}}_1$ is closest to $\mathcal{G}_1$, the middle section (k=i+3 to k=j+2) is closest to $\mathcal{G}_2$ and the final section (k=j+3 to k=K) is closest to $\mathcal{G}_3$. The second measure computed, θ_q^k, is the angle between the tangent to the qth zero track of $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$ and the tangent to the zero track of $Z\{\mathcal{G}(\zeta,\gamma)\}$ with the smallest $d_{pq}[k]$. This second error measure should provide an indication as to whether the qth zero track of $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$ and the zero track of $Z\{\mathcal{G}(\zeta,\gamma)\}$ with the smallest $d_{pq}[k]$ head in the same direction at the current value of k. The θ_p^k's are stored in the second of the two reference arrays for the qth zero track of $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$.

To splice together the zero tracks to obtain the best estimate of $Z\{\mathcal{H}(\zeta,\gamma)\}$, the zero track of $Z\{\mathcal{G}(\zeta,\gamma)\}$ which matched the qth zero track of $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$ best overall, on the basis of the Euclidean distance, is selected. This track is referred to from here on as the base track. The algorithm scans the two reference arrays for the qth zero track of $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$ and performs the following operations dependent on their contents at the current value of k:

- If the contents of the first reference array indicates that the base track does not contain the closest zero and the second reference array states that $\theta_p^k < \pi/4$, then the kth zero from the base track and the closest zero are swapped since it is likely the tracks have joined (splicing occurs). This is illustrated in Fig. 5.3 for k=i+5 to k=j and k=j+5 to k=K.

- If the contents of the first reference array indicate that the base track contains the closest zero to the psf's zero at the current value of k and the second reference array states that

$\theta_p^k < \pi/4$, then no zeros are swapped since the base track contains the zero which is best matched to the psf zero. This is illustrated in Fig. 5.3 for k=1 to k=i.

- If the second reference array states that $\theta_p^k > \pi/4$, then the closest zero and the kth zero in the base track are marked for removal (denoted by the symbol • in Fig 5.3). This action is taken because the track containing the closest zero is not heading in the same direction as the current psf track and the algorithm cannot determine, with a reasonable certainty, whether the track belongs to the psf or image. This is illustrated in Fig. 5.3 for k=i+1 to k=i+4 and k=j+1 to k=j+4.

Although the threshold for the angle between the tangents ($\pi/4$) is arbitrary, it is sufficiently large that inaccuracies in the psf estimate can be tolerated and sufficiently small that tracks clearly heading in different directions can be detected.

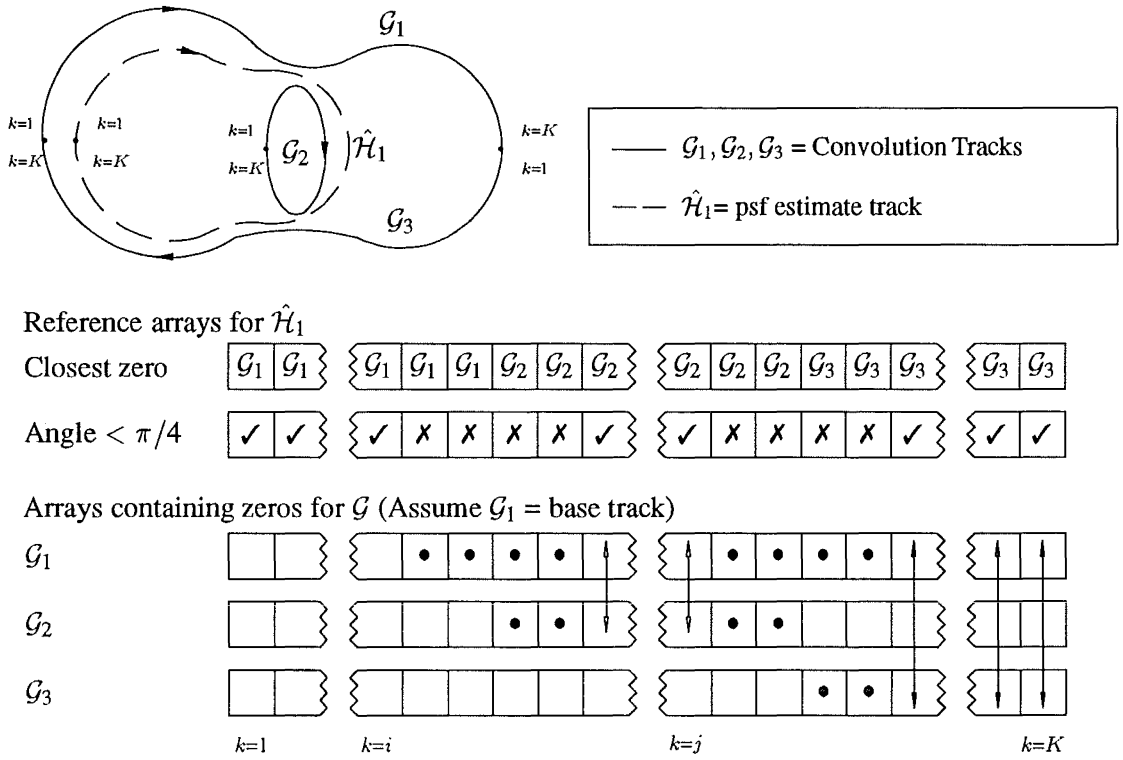

Figure 5.3 The convolution tracks $\mathcal{G}_1$,$\mathcal{G}_2$, $\mathcal{G}_3$ have been joined together as the blurred image has been contaminated with additive noise. The algorithm splices sections of the convolution tracks together, as explained in the text, to obtain the best possible approximation to $\mathcal{H}_1$. Swapping is indicated by arrows ↕, removal is indicated by •.

After the tracks have been spliced together the resulting psf and image tracks still in some cases contain small sections of track from the other component of the blurred image. It has been the author's experience that when zero tracks join together it is usual for significant sections of both sets of tracks to become joined together. If small sections of track are spliced into either the image or psf estimate tracks it most likely that a track that momentarily passes alongside a psf track has been incorrectly spliced-in. Thus, any spliced-in section of track containing less than 5% of the K zeros comprising the track has the zeros in the spliced in section removed. While the 5% threshold was chosen arbitrarily, experimentation with a number of convolutions derived from 5×5 pixel random images and a 5×5 pixel psf confirmed that the threshold was

sufficient to remove the small sections of track whose origin could not be determined with certainty.

A demonstration of the ability of this algorithm to correctly splice together the merged zero tracks of $\mathcal{G}(\zeta,\gamma)$ to obtain a good estimate of the zero tracks of $\mathcal{F}(\zeta,\gamma)$ is shown in Fig. 5.4. The zero tracks, for the slice corresponding to ρ = 1.05, of: a random 5×5 image, the psf in Fig. 5.2 (c), the 60dB SNR contaminated convolution formed from this psf and image, the image estimate zero tracks obtained with the basic Euclidean recognition scheme, and the image estimate zero tracks obtained from the recognition and splicing scheme are shown in Fig. 5.4 (a), (b), (c), (d), and (e) respectively. On comparing Figs. 5.4 (d) and (e) to Figs. 5.4 (a) and (b) it can be seen that the splicing scheme allows the merged tracks within the convolution's zero sheet to be separated, which in turn provides a set of image estimate tracks which are ideally uncontaminated by the psf tracks. It is worth bringing to the reader's attention that in Fig. 5.4 (d) the zero tracks pertaining to the psf can be clearly seen in the image estimate zero tracks provided by the basic Euclidean technique. The slight asymmetry in Fig. 5.4 (e) occurs because these asymmetrical portions of the image's zero tracks were originally part of the zero tracks that belong to the psf (i.e. these portions of image zero track had merged with the psf zero tracks). The portion of zero track that would have to be added to Fig. 5.4 (e) to preserve the symmetry was removed by the splicing algorithm, because it had a length less than 5% of the total track length.

COMPARISON OF ALGORITHMS INCORPORATING ZERO TRACK SPLICING AND NO SPLICING

	60 dB SNR $g(x,y)$			40 dB SNR $g(x,y)$		
ρ	Q dB No Splice	Q dB Splice	ΔQ	Q dB No Splice	Q dB Splice	ΔQ
0.6	24.0	23.4	-0.6	6.1	2.3	-3.8
0.7	29.5	30.7	1.2	8.0	8.1	0.1
0.8	32.8	33.5	0.7	2.1	12.4	10.3
0.9	2.7	8.8	6.1	3.5	4.9	1.4
1.0	2.4	29.0	26.6	0.6	6.3	5.7
1.1	35.4	36.6	1.2	3.9	3.9	0.0

Table 5.2 Key: ΔQ = Q dB Splice - Q dB No Splice.

The same set of 60 dB and 40 dB SNR convolutions used to assess the performance of the basic Euclidean technique were deconvolved using a version of the consistent algorithm which incorporated this new recognition and separation scheme. The results from this investigation, which are presented in Table 5.2 under the heading 'Splice', clearly show that the scheme using zero track splicing is in the majority of cases able to dramatically improve the quality of the reconstructions which previously suffered from the effects of zero sheet merging.

5.3 RECONSTRUCTING THE IMAGE

Once the estimate of $Z\{\mathcal{F}(\zeta,\gamma)\}$ has been obtained, $\hat{f}(x,y)$ is acquired, from samples of $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$, by using the algebraic reconstruction technique. The motivation for using the algebraic reconstruction technique was that the required samples can be selected from anywhere on the zero sheet, whereas with the Fourier based method there are tight constraints on the position of the samples (see §4.5.3). It is desirable to have a reasonable amount of flexibility

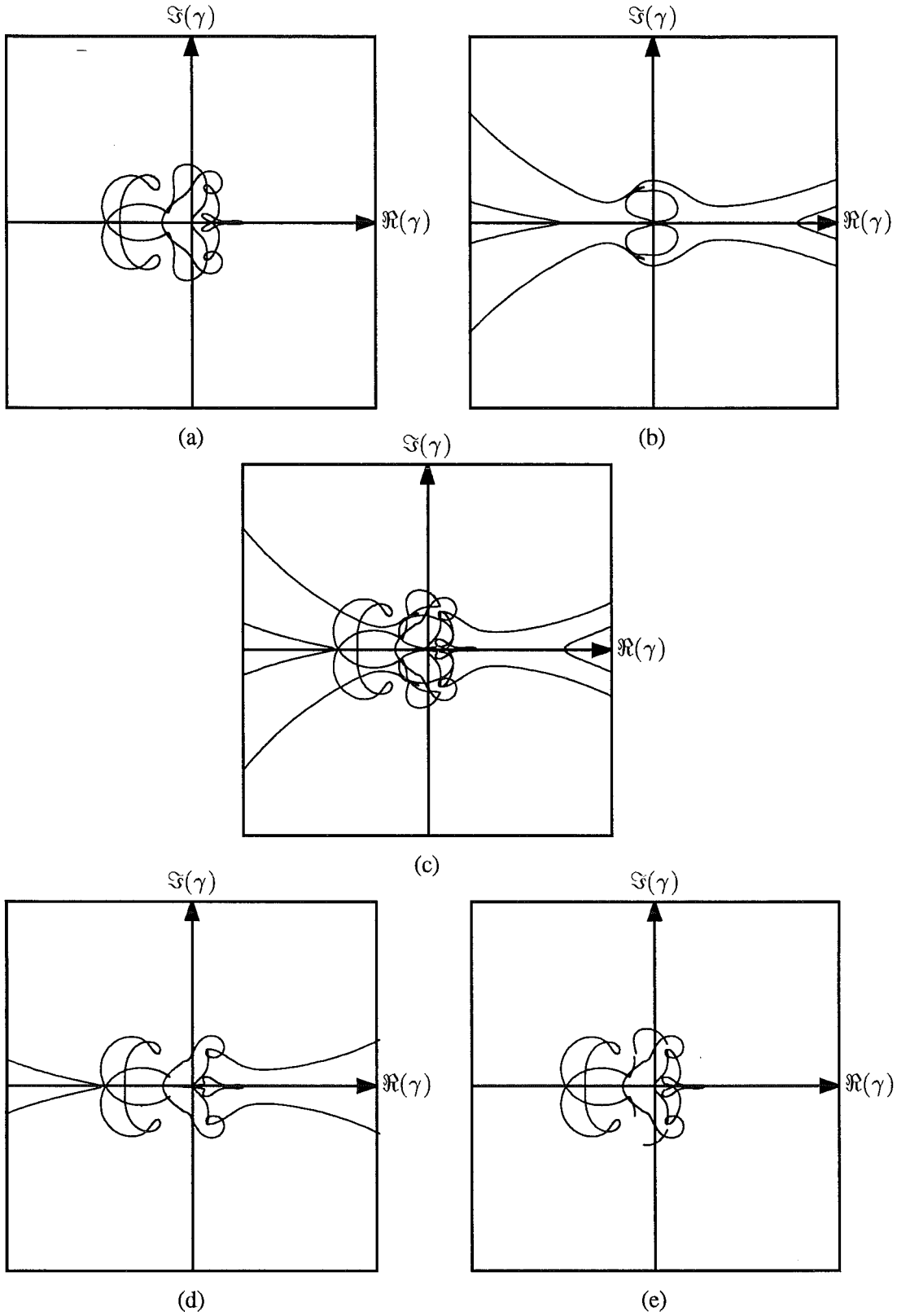

Figure 5.4 The zero tracks of: (a) a random 5×5 image, (b) the 5×5 psf in Fig. 5.2 (c), (c) the 60dB SNR contaminated convolution formed from the image and psf in (a) and (b), (d) the image estimate zero tracks obtained with the basic Euclidean recognition scheme, and (e) the image estimate zero tracks obtained from the recognition and splicing scheme. The slices correspond to $\rho = 1.05$

over the selection of the zero samples as contamination, in some situations, distorts some zeros more than others (see §4.4.2) and it may not always be possible to classify all the zeros comprising $Z\{\mathcal{G}(\zeta,\gamma)\}$ as belonging to the psf or image. The flexibility of the algebraic reconstruction technique, with respect to the sampling of the zeros, prompts the question: where from the zero sheet should the zero samples be taken to guarantee that the best possible image estimate is obtained? The author has considered this question and developed a sampling technique which is able to reliably provide good quality reconstructions. In this section the principles of the sampling technique are described and results demonstrating the performance of the algorithm are presented. For reasons of clarity the sampling technique that is used in the situation where only one slice from $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$ and $Z\{\mathcal{G}(\zeta,\gamma)\}$ is used to deconvolve the image is described first. This is followed by discussions which describe the extension of the sampling technique to the situation where several slices from $Z\{\hat{\mathcal{H}}(\zeta,\gamma)\}$ and $Z\{\mathcal{G}(\zeta,\gamma)\}$ are used to perform the deconvolution and the situation where it is known *a priori* that the contamination is band-limited.

5.3.1 Reconstruction from a Single Slice

A technique which selects zero samples from a slice of $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ and constructs the set of equations required by the algebraic reconstruction technique has been developed. The research and development process that ultimately resulted in this sampling technique was driven by the need to produce a sampling technique that could sample $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ in a manner that provided the best possible $\hat{f}(x,y)$ in the presence of wide-band contamination. As noted in §4.4.1 in the presence of wide-band contamination it is not possible to know which of the zeros comprising $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ have been perturbed most by the contamination and all zeros must therefore be considered to have been perturbed by the same amount. For the purposes of this discussion the best possible $\hat{f}(x,y)$ is the $\hat{f}(x,y)$ with the highest Q that can be obtained from the slice of $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ provided by the zero sheet recognition and separation scheme described in §5.2. The best possible $\hat{f}(x,y)$ is most likely to be that obtained from those samples of $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ that provide the most well-conditioned set of equations for the algebraic reconstruction technique.

A well-conditioned set of equations is a set of equations which is linearly independent (Westlake, 1968). To ensure that the set of equations formed from the samples of $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ are as well conditioned as possible, in the presence of any amount of contamination, it is necessary to sample zeros from $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ which result in a linearly independent set of equations. Zero samples that are immediately adjacent in Z space are going to give rise to equations which are likely to be linearly dependent. This property therefore suggests that the most straightforward method that could be used to select zero samples, which give rise to a linearly independent set of equations, is to select zeros that are widely separated in Z space. Thus the specification for the sampling technique developed was that it must be able to quickly produce a set of zero samples from the zero sheet which are maximally displaced in the 4-D space in which the zero sheet exits. In the situation where the contamination is band-limited it is likely that a well-conditioned set of equations would result if samples were predominantly selected from those regions of the zero sheet known to be least effected by the contamination. A sampling method for images that have been contaminated by band-limited contamination is presented in §5.3.3.

Developing an algorithm that can sample a slice of $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ to produce a maximally displaced set of zero samples is not a simple task. The reason for this is that the properties of the zero sheet described in §4.1.1 indicate that it is not possible to use a technique that simply

selects zeros that are maximally displaced in either the ζ or γ planes, because the corresponding zeros in the γ or ζ planes tend to be very close together. As examples consider the two sampling schemes described below which are possibly the most simple sampling schemes that could be used. The first sampling scheme involves choosing a slice corresponding to a circular contour that has a large extent in the ζ plane (i.e. ρ in Eqn.(4.6) is large) and choosing points around this contour which are equally spaced. The problem with this sampling technique is that while the points are maximally displaced in the ζ plane the corresponding zeros in the γ plane are going to be very close together, because as $\rho \rightarrow \infty$ the zero sheet slices in the γ plane shrink to a set of point zeros. It is also worth noting that sampling evenly around the circular contour also has the major limitation that, in the contaminated situation, the corresponding zero samples in the γ plane may not always be available (due to an inability to characterise particular γ plane zeros as belonging to the psf or image).

The second simplistic technique involves finding the slice that has the maximum extent in the γ plane and then sampling zeros that are maximally displaced in this slice. The difficulty with this technique is that while zero samples of large magnitude on the same zero track are quite widely separated in the γ plane they correspond to points quite close together on the corresponding circular contour in the ζ plane (see §4.2.1). Later in this section results are presented which confirm that these simple sampling techniques generate sets of equations with less than optimal conditioning and $\hat{f}(x, y)$ of less than optimal Q.

The sampling technique developed by the author is a straightforward sampling technique which attempts to find zero samples that are maximally displaced in the γ plane while avoiding sampling from those regions of the zero sheet, outlined above, that are likely to correspond to points which are close together in the ζ plane. The technique also takes advantage of the property of the QR decomposition algorithm that an estimate of the condition number, $\kappa(\mathbf{A})$, of the system of equations can be obtained from the ratio $\mathbf{R}_{00}/\mathbf{R}_{PP}$ as explained in §4.5.3. In the situation where a slice is to be used to deconvolve the image, the author's sampling technique requires a value for ρ to be chosen that produces a slice which has a large extent in the γ plane. It is not difficult to select a suitable value for ρ because experience has shown that the slices with the largest extent usually exist for ρ values of around 1.0. By manually selecting a ρ value around 1.0 the problems associated with choosing a ρ value which is so large that zeros comprising the slice are too close together is avoided. Further on in this section experimental results that confirm that the best quality reconstructions are obtained from slices with ρ values of approximately 1.0 are presented.

On obtaining the specified $Z\{\hat{\mathcal{F}}(\zeta, \gamma)\}$ slice from the recognition and separation algorithm the sampling algorithm first computes the total Euclidean length of the sections of zero track, comprising the $Z\{\hat{\mathcal{F}}(\zeta, \gamma)\}$ slice, that lie within some preset radius $|\gamma| < r_\gamma$. The total zero track length which lies within the radius r_γ is divided by the total number of samples required, to give the distance along each section of track within r_γ at which a zero sample is to be taken. By setting r_γ to a suitable value, those zeros that are widely separated in γ space (which tend to have large magnitudes) and which have corresponding points in ζ space that are close together should not be selected. If the condition number for the resulting set of equations is computed and found to be unsatisfactory, r_γ can be changed to give a new set of zero samples and hopefully a better conditioned set of equations. Although the set of zero samples obtained using this technique cannot be said to be maximally separated, provided ρ and r_γ are chosen appropriately, the distribution of samples is generally wide enough to obtain a well-conditioned set of equations. Experimental results presented later confirm that if r_γ is chosen appropriately a well-conditioned set of equations, which produce a good quality $\hat{f}(x, y)$, can be obtained from this sampling method.

As explained in §4.1.2, $2(N-1)^2+1$ zero samples are required to guarantee a unique reconstruction, however the author uses $2N^2+1$ zero samples, as experience has shown that using more than the minimum number of zero samples tends to make the reconstruction process more reliable. Evidence of the improved reliability of using more than the minimum number of zero samples is shown in Fig. 5.5 where the condition number κ (see §4.5.3), of the system of equations defining the image, is plotted against r_γ for $2(N-1)^2+1$ and $2N^2+1$ zero samples. Since a smaller value of κ indicates a better conditioned set of equations, it can be seen that using $2N^2+1$ zero samples almost always results in a set of equations of better conditioning than if $2(N-1)^2+1$ zero samples are used. It can also be seen that increasing the number of zero samples also tends to make the κ *vs* r_γ plots smoother, indicating that using more samples makes the reconstruction process less sensitive to the choice of r_γ. The above finding is in agreement with Curtis and Oppenheim (1987) and Zakhor and Oppenheim (1990) who have also observed that using more than the minimum number of zero samples improves the reliability of the reconstruction process.

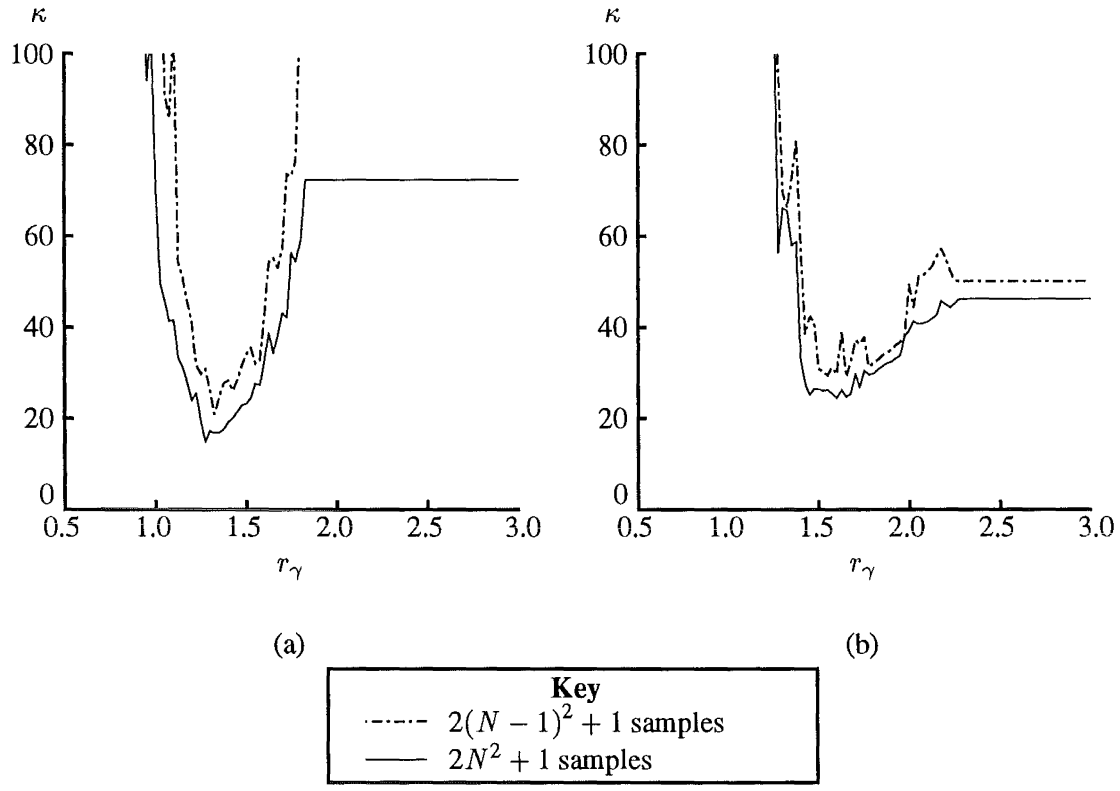

Figure 5.5 Evidence that using more zero samples than the minimum number $(2(N-1)^2+1)$ required to obtain a reconstruction, improves the conditioning of the set of equations that define the image estimate. (a) and (b) are plots of the condition number, κ, *vs* r_γ for two different 5×5 random images.

To determine which value of r_γ is likely to provide an $\hat{f}(x,y)$ with the best Q, the author used a version of the above sampling technique which varied r_γ between 0.5 and 3.0 in 100 equal increments. Image estimates were obtained from each set of zero samples that were selected with each increment in r_γ. A comprehensive study of the dependence of Q and κ on r_γ has been conducted using the above sampling technique and a number of blurred images of varying sizes. The images used were various random 5×5, 8×8, and 16×16 pixel images convolved with the bilevel 5×5 pixel psf in Fig. 5.2 (c). The convolutions presented to the consistent algorithm were either uncontaminated or contaminated with SNRs of 60dB or 40dB.

At present the consistent algorithm is unable to obtain useful reconstructions from original images larger than 16×16 pixels in the situation where the convolution has a 40 dB SNR and for this reason the study was limited to images of an extent $\leq$ 16×16. In the following discussion only results from the 16×16 images are presented as these results, for the most part, have been found to be representative of those obtained from the smaller images. If any dependence on image size was noted this is explicitly stated. To determine how Q and κ depend on r_γ, the values of Q and κ obtained from the above set of images were plotted against r_γ for the range of r_γ values specified.

In Fig. 5.6 representative Q *vs* r_γ and κ *vs* r_γ plots obtained from uncontaminated (Fig. 5.6 (a) and (b)) and contaminated (Fig. 5.6 (c) and (d)) versions of two different 20×20 convolutions (16×16 original images) are shown. The $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ slice from which the zero samples were selected to produce Figs. 5.6 (a) and (b) and Figs. 5.6 (c) and (d) were those corresponding to $\rho = 0.8$ and 1.0 respectively. The study showed (cf. Fig. 5.6):

- The best quality $\hat{f}(x,y)$ (i.e. that with the largest Q) is obtained from those samples which result in a set of equations that has a κ that is amongst the smallest. This indicates that κ can indeed be used to indicate which samples are likely to give the best $\hat{f}(x,y)$.
- The best quality $\hat{f}(x,y)$ and the set of equations with the smallest κ are usually obtained from zero samples that have been produced by setting r_γ in the range 1.0 to 1.5.
- If r_γ is made too small, the samples in γ space are too close together and as a consequence the resulting set of equations are ill-conditioned and $\hat{f}(x,y)$ is of a poor quality.
- As r_γ becomes too large (i.e. as $r_\gamma \rightarrow 3.0$) and the samples become equispaced along the entire length of the zero tracks forming the $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ slice, the set of equations produced from the set of samples becomes ill-conditioned, i.e κ increases. This occurs because as $r_\gamma \rightarrow \infty$ the corresponding points in the ζ plane become closer and closer together, reducing the linear independence of the set of equations and making them ill-conditioned.

In addition to the above findings, it was found that the minimum value of r_γ that resulted in a useful image estimate was dependent on the image size. This was to be expected because as the image size increases the number of samples used to reconstruct the image increases and as a consequence the zero samples must become closer together in the γ plane. In summary the study has shown, for images of the size range used above, that the optimum value of r_γ is between 1.0 and 1.5 and that κ can be used to indicate which value of r_γ has produced the samples that are likely to give the best $\hat{f}(x,y)$.

A sampling algorithm that varies r_γ in small increments between 1.0 and 1.5 and uses κ to select the best set of zero samples is impractical due to the large amount of computation required to compute κ. In an attempt to develop a practical sampling scheme the author adapted the sampling scheme so that it only obtained zero samples and image estimates for three different values of r_γ between 1.0 and 1.5. The algorithm then chooses the set of zero samples corresponding to the smallest κ to form the final image estimate, since this image estimate should be the best of the three possible estimates. The sets of three r_γ values which the author choose to use were either 1.0, 1.25, and 1.5 or 1.125, 1.25, and 1.375. To evaluate the performance of this adapted sampling scheme, the set of images described earlier were deconvolved by a consistent algorithm incorporating this sampling scheme. The quality of the $\hat{f}(x,y)$ obtained from this adapted sampling technique were compared to those obtained

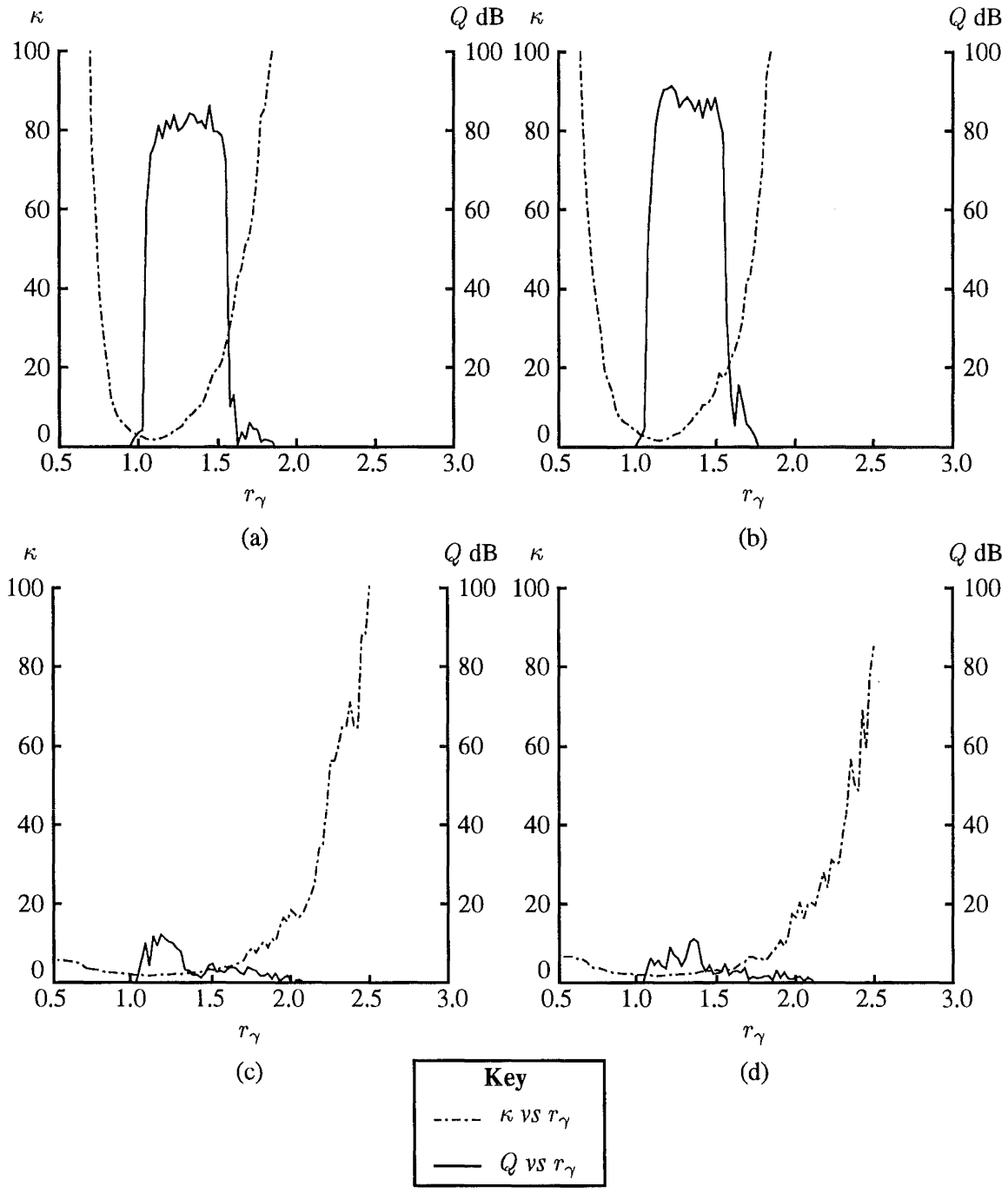

Figure 5.6 Representative quality factor (Q) *vs* r_γ and condition number (κ) *vs* r_γ plots obtained from uncontaminated ((a) and (b)) and contaminated ((c) and (d)) versions of two different 20×20 convolutions.

from a technique that varied r_γ between 1.0 and 1.5 in increments of 0.025 and used prior knowledge of $f(x,y)$ to select the best image estimate from all the reconstructions. In the following discussion the sampling technique using three r_γ values is known as the 'coarse' r_γ technique while the second technique is known as the 'fine' r_γ technique.

In Table 5.3 representative results comparing the quality of the $\hat{f}(x,y)$ obtained from the coarse r_γ technique to those obtained from the fine r_γ technique are presented. The results in Table 5.3 were produced from the same 60 dB SNR 20×20 pixel convolutions used to produce Fig. 5.6 and the zero sheet slices used were produced for ρ values ranging from 0.5 to 1.2. The comparison shows that the coarse r_γ technique is unable to obtain reconstructions of the same quality as the fine r_γ technique and the difference in the quality of the reconstructions is quite large (i.e. up to 4 dB). The quality of the best $\hat{f}(x,y)$ obtained from the coarse r_γ technique are, however, of a useful accuracy and as a consequence the author believes that this technique does hold some promise. The author has also experimented with taking the mean of the three $\hat{f}(x,y)$ obtained from the coarse r_γ technique to determine whether this method could improve the $\hat{f}(x,y)$. It was found that while the mean of the three $\hat{f}(x,y)$ was generally not of the same quality as those obtained from the standard coarse r_γ technique, the averaging process did have a stabilising effect on the quality of the reconstructions. A major difficulty with computing the mean image estimate in practice is that the three $\hat{f}(x,y)$ are all scaled by different complex constants and a technique has to be developed to equate all the complex constants. It should be possible to develop a least squares technique that scales all the image estimates to within the same complex constant and as a result this technique which averages the three image estimates is probably also worthy of further consideration. All of the results presented in the remainder of this thesis have been produced using the fine r_γ technique unless otherwise specified.

COMPARISON OF COARSE AND FINE r_γ ALGORITHMS

	20×20 $g_1(x,y)$			20×20 $g_2(x,y)$		
ρ	Q dB Fine	Q dB Coarse	ΔQ	Q dB Fine	Q dB Coarse	ΔQ
0.5	0.00	0.00	0.00	0.02	0.00	0.02
0.6	0.00	0.00	0.00	0.03	0.00	0.03
0.7	1.17	0.00	1.17	0.06	0.00	0.06
0.8	3.26	0.05	3.21	2.15	0.22	1.93
0.9	10.11	7.16	2.95	10.70	8.90	1.80
1.0	12.23	8.25	3.98	11.88	8.34	3.54
1.1	5.67	4.69	0.98	6.20	6.20	0.00
1.2	2.59	0.98	1.61	3.41	2.60	0.81

Table 5.3 Key: ΔQ = Q dB Fine - Q dB Coarse. The subscripts denote the convolutions derived from different original 16×16 pixel images.

The final factor which has an effect on the quality of the $\hat{f}(x,y)$ obtained from the sampling technique described here is the choice of ρ. It was stated earlier that $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ slices for ρ values around 1.0 have the largest extent in the γ plane and so the sampling algorithm is able to sample zeros which are maximally displaced in Z space. To confirm that the best reconstructions are obtained from the slices corresponding to $\rho \approx 1$, the consistent algorithm was used to convolve the set of images described previously by using zero sheet slices that corresponded to various ρ values. In this evaluation the $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ slices corresponding to ρ values of 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, and 1.2 were produced for each image and the quality of the reconstructions obtained from each slice were compared.

In Fig. 5.7 a representative graph showing the dependence of Q on the selected value of ρ are shown. These graphs are produced from the same uncontaminated and contaminated 20×20 pixel convolutions used to produce Fig. 5.6. The results from this evaluation demonstrated, and the graphs in Fig. 5.7 confirm, that the best reconstructions are obtained from zero sheet slices corresponding to ρ values around 1.0. The graphs in Fig. 5.7 clearly show a maximum around $\rho = 1$ with a noticeable drop in performance for $\rho < 0.9$ and $\rho > 1.1$. These results also confirm that the shrinking of the zero sheet as $\rho \rightarrow 0$ or $\rho \rightarrow \infty$ forces the zero samples to be closer together, producing ill-conditioned sets of equations (i.e. κ increases) and poor quality $\hat{f}(x, y)$. The drop in performance as $\rho \rightarrow \infty$ also confirms the notion described earlier that choosing points around a circular contour with a large ρ value would produce poor quality image estimates, due to the shrinking of the zero sheet and the corresponding minimal displacement of the corresponding zero samples in the γ plane.

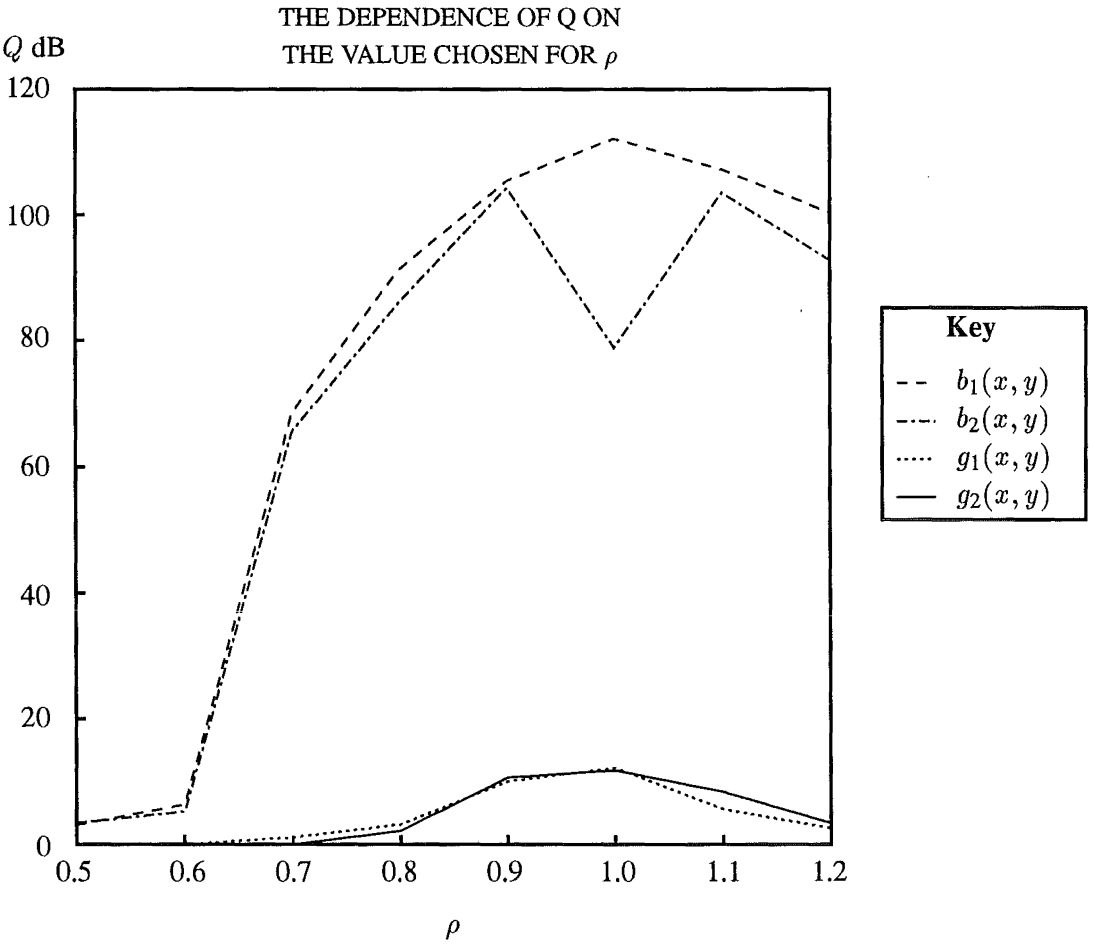

Figure 5.7 The dependence of Q on the selected value of ρ for two 20×20 convolutions (original images 16×16) with (g(x,y)) and without (b(x,y)) contamination (60 dB SNR).

5.3.2 Reconstruction from Several Slices

It was anticipated that an improvement in the quality of the $\hat{f}(x, y)$ obtained from the consistent algorithm could be obtained if the recognition, separation and sampling of $Z\{\hat{\mathcal{F}}(\zeta, \gamma)\}$ was performed using more than one slice. The primary motivation for this belief was that by using more than one slice it should be possible to displace the zero samples further in Z space,

thereby obtaining sets of equation with improved conditioning and better $\hat{f}(x,y)$. The second property which should enable better $\hat{f}(x,y)$ to be obtained, is that it should be possible to include zero samples from the slices of $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ for very large or very small ρ values which tend to be very accurate. These slices cannot be used to produce a reconstruction on their own as the extent of the zero tracks at these values of ρ is very small, however the decline in the likelihood of zero sheet merger at these values of ρ (see §4.4.3) results in $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ slices which tend to be very accurate.

A new version of the consistent algorithm that could recognise and separate $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ in several slices was implemented. This algorithm is henceforth referred to as the multiple slice algorithm. The sampling technique was also adapted so that it summed over all the slices the total Euclidean length of the sections of zero track, comprising the $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ slices, that were within r_γ. This total track length, as in §5.3.1, was then divided by the total number of samples required to give the distance along each track at which a sample was to be taken. This new sampling algorithm therefore allocates more samples to those slices that have a larger extent in the γ plane since these tracks contribute most to the track length summation, i.e. as ρ increases from 0 to 1.0 the number of samples taken from the corresponding slices increases.

To confirm that it is beneficial to use more than one slice of $Z\{\mathcal{G}(\zeta,\gamma)\}$ to obtain a reconstruction, an evaluation of the performance of the multiple slice algorithm was carried out. This evaluation was conducted using the set of contaminated and uncontaminated images described in §5.3.1 and the number of slices used ranged from one to five. The slices were produced for ρ values ranging from 0.5 to 1.2 and in the multiple slice situation the increment in the ρ value of successive slices was 0.1. In Table 5.4 representative results, from the 60 dB SNR and 40 dB SNR 20×20 pixel convolutions, for varying numbers of slices, are presented. The results in Table 5.4 clearly show that by using more than one slice it is possible to obtain an $\hat{f}(x,y)$ of a quality superior to that which can be obtained by using a single slice (e.g. for the 60 dB SNR $g_1(x,y)$, the reconstruction from one slice had a Q of 12.23 dB while that from five slices had a Q of 19.10 dB). It is necessary to note, especially for the lower SNR convolutions, that there is no clear relationship between Q and the number of slices, as the number of slices is increased. In fact the 40 dB SNR results in Table 5.4 show that Q will increase with the number of slices (this is due to the greater separation of the zero samples resulting in a set of equations with greater linear independence) until a slice, which is more inaccurate than the preceding slices, is included. This inaccurate slice produces a corresponding decrease in Q. The author also noted that using more slices had a stabilising effect on the consistent algorithm as the quality of the image estimates obtained from more than one slice tend to be less sensitive to the choice of r_γ and the effects of contamination. This increase in robustness of the algorithm appears to be a function of the number of slices used as the sensitivity of the reconstruction Q to the above factors seemed to decrease as the number of slices increased.

5.3.3 Weighted QR Decomposition and Band-limited Contamination

In the situation where prior knowledge indicates that some zeros comprising $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ are likely to be more accurate than others, for the purposes of image reconstruction, it would be desirable to sample the more accurate zeros in preference to the other zeros. While in some cases it may be possible to sample only those zeros from the regions of the zero sheet where the zeros are known to be accurate, it is more likely that, if the resulting set of equations are to be linearly independent, zeros are going to have to be also sampled from the inaccurate regions. What is required therefore is a technique that allows the more accurate zero samples to have more influence on the form of the $\hat{f}(x,y)$ than the less accurate samples. A technique

MULTIPLE SLICES AND THE IMPROVEMENT IN IMAGE ESTIMATE QUALITY

Slices	Q dB 60 dB $g_1(x,y)$	Q dB 60 dB $g_2(x,y)$	Q dB 40 dB $g_1(x,y)$	Q dB 40 dB $g_2(x,y)$
1	12.23 (1.0)	11.88 (1.0)	2.04 (0.9)	4.00 (1.1)
2	16.61 (0.9)	13.44 (0.9)	3.42 (0.9)	4.28 (0.9)
3	16.95 (0.8)	15.17 (0.9)	4.01 (0.8)	4.41 (0.8)
4	17.45 (0.8)	14.50 (0.7)	3.93 (0.7)	4.74 (0.8)
5	19.09 (0.6)	17.63 (0.7)	3.79 (0.6)	4.04 (0.7)

Table 5.4 The subscripts denote the convolutions derived from different original 16×16 pixel images. The Qs in this table are the maximum of all the Qs obtained for the indicated image using the indicated number of slices. The numbers in brackets indicate the value of ρ that corresponds to the first slice in the series of slices used to obtain the $\hat{f}(x,y)$ with the maximum Q.

which allows this principle to be implemented is the weighted QR decomposition technique, described in §4.5.2.2, which allows the equations derived from the more accurate samples to be weighted so that their residuals are smaller than the residuals associated with the inaccurate equations. Examples of situations where weighted QR decomposition could be used to improve the quality of the image estimates obtained, are the situations where it is known that some zeros comprising $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ are more likely to have been involved in zero sheet merging than others and when some zeros have been affected more by the contamination than others. The situation where contamination effects the zero sheet more in some regions than others is dealt with here.

It was shown in §4.4.2 that in the situation where an image is corrupted by band-limited contamination, some regions of the zero sheet are perturbed more by contamination than other regions. If prior knowledge is available on the bandwidth of the band-limited contamination it is possible to predict which regions of the zero sheet are most likely to have been affected by the contamination. It should therefore be possible to weight heavily the equations produced from the samples from the remaining regions of the zero sheet and obtain an $\hat{f}(x,y)$ Q which is superior to that which would otherwise be obtained. To determine whether in practice the above technique could be used to improve the quality of the $\hat{f}(x,y)$, obtained in the situation where the contamination is band-limited, a study has been conducted. In this study a series of 8×8 and 16×16 random images were convolved with the 5×5 bilevel psf in Fig. 5.2 (c) and pseudo random noise with various cut-off frequencies was added to the resulting convolutions. The resulting convolutions were then deconvolved using a version of the consistent algorithm which incorporated a weighted QR decomposition scheme that used various weighting factors.

In Table 5.5 a representative example of the results obtained from this investigation is given. The results in Table 5.5 were obtained for the case where either high frequency or low frequency contamination with a cut-off frequency of $BW_c^l = \frac{1}{4\epsilon}$ or $BW_c^h = \frac{1}{4\epsilon}$ was applied to $B(u,v)$ respectively (see Fig. 5.8) As in §4.4.2 the superscripts l and h denote the lowest and highest frequency component affected by the contamination respectively. The $\hat{f}(x,y)$ were obtained using a weighted QR decomposition scheme, which in the high frequency case weighted equations produced from the zero samples within $BW_w^h = \frac{1}{8\epsilon}$, $BW_w^h = \frac{1}{4\epsilon}$, $BW_w^h = \frac{3}{8\epsilon}$, or $BW_w^h = \frac{1}{2\epsilon}$ more heavily than those produced from zero samples taken from the remaining regions of the spectrum and in the low frequency case weighted those equations from zero samples within $BW_w^l = \frac{1}{8\epsilon}$, $BW_w^l = \frac{1}{4\epsilon}$, $BW_w^l = \frac{3}{8\epsilon}$, or $BW_w^l = 0$ heavier than the remaining equations (see Fig. 5.8). When $BW_w^h = \frac{1}{2\epsilon}$ or $BW_w^l = 0$ the entire spectrum of $G(u,v)$ is weighted uniformly and therefore these two cases correspond to the unweighted situation. The two weighting ratios used with the weighted QR decomposition scheme were 4:5 and 1:2.

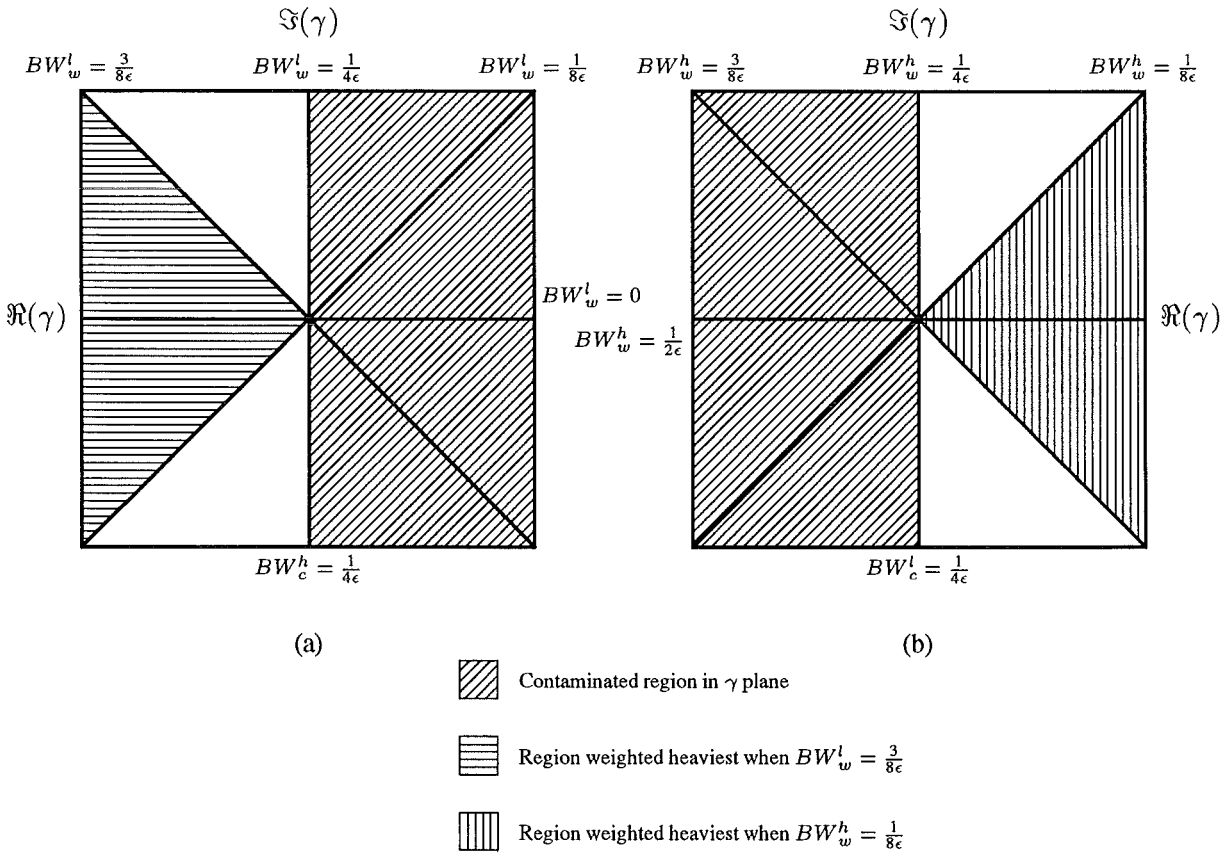

Figure 5.8 The regions of Z space that were corrupted by band-limited contamination to produce the results in Table 5.5. Shown in (a) is the low frequency contamination case, where the cut-off frequency was $BW_c^h = \frac{1}{4\epsilon}$ and so the left hand half of the γ plane is contaminated. In (b) the high frequency contamination case is shown, with the right hand half of the γ plane contaminated as the contamination cut-off frequency was $BW_c^h = \frac{1}{4\epsilon}$. Also shown in (a) and (b) are the respective regions in γ space where the zeros are weighted heaviest. For example, in (a) when $BW_w^l = \frac{3}{8\epsilon}$ the zeros in the shaded region indicated, are weighted heavier than those in the remaining region. As BW_w^l decreases, the shaded region rotates from $BW_w^l = \frac{3}{8\epsilon}$ around to $\frac{1}{4\epsilon}$, $\frac{1}{8\epsilon}$, and finally to 0, where the entire γ plane is weighted by the heaviest weight. The weighting shown in (b) operates in a similar manner.

In the situation where the contamination was confined to low frequencies less than $\frac{1}{4\epsilon}$, it can be seen in Table 5.5 that the best image estimates were obtained when those equations derived from zero samples corresponding to $|\mathbf{u}| \geq \frac{1}{4\epsilon}$ (i.e $BW_w^l = \frac{1}{4\epsilon}$ and $BW_w^l = \frac{3}{8\epsilon}$) were weighted heavier than those equations derived from zero samples corresponding to $|\mathbf{u}| < \frac{1}{4\epsilon}$. As an example, for the 60 dB SNR low frequency contaminated convolution the weighted reconstruction corresponding to $BW_w^l = \frac{1}{4\epsilon}$ had a Q of 27.03 dB, when the weighting ratio was 4:5, while the unweighted reconstruction ($BW_w^l = 0$) had a Q of 26.84 dB. In the situation where the contamination was confined to frequencies greater than $\frac{1}{4\epsilon}$, it can be seen that the best image estimates were generally obtained when those equations derived from the zero samples corresponding to $\mathbf{u} \leq \frac{1}{4\epsilon}$ (i.e $BW_w^h = \frac{1}{4\epsilon}$ and $BW_w^h = \frac{1}{8\epsilon}$) were weighted heavier than those equations produced from other samples. To illustrate the above result consider the 60 dB SNR high frequency contaminated convolution which provided an image estimate with a Q of 18.26 dB in the weighted case, where $BW_w^h = \frac{1}{4\epsilon}$ and the weighting ratio was 4:5, and compare this to the inferior $\hat{f}(x, y)$, with a Q of 17.76 dB, provided in the unweighted case

RECONSTRUCTION USING WEIGHTED QR DECOMPOSITION
WHEN THE CONVOLUTION CONTAMINATION IS BAND-LIMITED

Low Frequency Contamination			
SNR	BW_w^l	Q dB with WR = 4:5	Q dB with WR = 1:2
60dB	$\frac{3}{8\epsilon}$	26.95 *	27.10 *
	$\frac{1}{4\epsilon}$	27.03 *	28.06 *
	$\frac{1}{8\epsilon}$	26.32	25.87
	0	26.84	26.84
40dB	$\frac{3}{8\epsilon}$	12.65 *	12.70 *
	$\frac{1}{4\epsilon}$	12.73 *	12.65 *
	$\frac{1}{8\epsilon}$	11.91	10.13
	0	12.54	12.54
High Frequency Contamination			
SNR	BW_w^h	Q dB with WR = 4:5	Q dB with WR = 1:2
60dB	$\frac{1}{8\epsilon}$	17.90 *	18.09 *
	$\frac{1}{4\epsilon}$	18.26 *	19.03 *
	$\frac{3}{8\epsilon}$	18.06 *	17.81 *
	$\frac{1}{2\epsilon}$	17.76	17.76
40dB	$\frac{1}{8\epsilon}$	1.59 *	1.69 *
	$\frac{1}{4\epsilon}$	1.40	1.26
	$\frac{3}{8\epsilon}$	0.86	0.00
	$\frac{1}{2\epsilon}$	1.46	1.46

Table 5.5 Key: WR = Weighting Ratio, * = weighted reconstruction with an accuracy superior to the unweighted reconstruction. In the low frequency contamination case the contamination was confined to frequencies $> |\frac{1}{4\epsilon}|$ and in the high frequency case was confined to frequencies $< |\frac{1}{4\epsilon}|$. BW_w^l indicates that all zero samples with frequencies greater than the specified value were weighted, while BW_w^h indicates that all zero samples with frequencies less than the specified value were weighted. In the low frequency case the unweighted results are presented in the row of the table corresponding to $BW_w^l = 0$ (i.e. uniform weighting) and in the high frequency case the unweighted results are in the row corresponding to $BW_w^h = \frac{1}{2\epsilon}$.

($BW_w^h = \frac{1}{2\epsilon}$). The results in Table 5.5 therefore confirm that the weighted QR decomposition scheme can be used to improve the quality of the $\hat{f}(x,y)$ obtained in the situation where the convolution is contaminated with band-limited contamination and prior knowledge of the contamination bandwidth is available. It should be noted however that this technique cannot be used to dramatically improve the quality of a poor $\hat{f}(x,y)$ as the improvements in Q are generally only of the order of a fraction of a decibel.

The image used in this study had the predominant low frequency nature which is a characteristic of real world images (see §3.2.2). In Table 5.5 it can be seen that the reconstructions obtained from the low frequency contaminated convolutions were always superior to those obtained from the high frequency contaminated convolutions. The author believes that the reason for this is that the low frequency contamination tends to distort the zero sheets of typical images less than high frequency contamination of the same energy, consequently a more accurate $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ and $\hat{f}(x,y)$ can be obtained in the low frequency situation. In §4.4.2 it was shown that low frequency contamination does indeed tend to distort the zero sheet of a typical image less than high frequency contamination of the same energy.

5.4 EFFECTS OF COMPONENT IMAGE SIZE AND CONTAMINATION LEVEL ON IMAGE ESTIMATES

The dependence of the quality of $\hat{f}(x, y)$ on the linear independence of the set of equations used to produce $\hat{f}(x, y)$ suggests that quality will also be affected by original image size and contamination level. An investigation of these effects, along with a study of the dependence of the reconstruction Q on the psf size, are presented here.

To determine whether the size of the original image could in some way influence the quality of the image estimate obtained from the consistent algorithm a study using original images of varying sizes has been conducted. The set of images used in this study were the set of 5×5, 8×8, and 16×16 random images convolved with a 5×5 psf, described in §5.3.1. This set of original images consists of two subsets of images comprising one 5×5, 8×8, and 16×16 pixel image each. Each of the subsets of images were derived from a random number generator which used a different seed for each subset of images. As a consequence the following set of results that compare image size to $\hat{f}(x, y)$ Q have been grouped according to the seed used to produce the particular subset of images (i.e. seed 1 or seed 2).

Presented in Table 5.6 are representative results obtained from the consistent algorithm when applied to either uncontaminated or contaminated blurred images of various sizes. The results demonstrate that as the size of the original image increases there is a corresponding drop in the quality of the reconstructions obtained from the consistent algorithm. The dependence of Q on image size can be explained by first acknowledging that an increase in image size, increases the number of zero samples that must be selected from within the region defined by r_γ, consequently the zero samples must move closer together and the linear independence of the set of equations must decrease. This decrease in the linear independence, or conditioning, of the set of equations results in the observed decrease in the $\hat{f}(x, y)$ Q as image size increases. In the contaminated situation there is an additional effect that decreases Q with increasing image size. This additional effect is the increased likelihood of zero track merger, as the number of tracks within r_γ increases with increasing image size. This increase in the likelihood of merger produces a corresponding decrease in the reconstruction Q as the algorithm now finds it more difficult to correctly classify all the zero tracks.

THE DEPENDENCE OF Q ON ORIGINAL IMAGE SIZE

Original Image Size	Uncontaminated $g(x, y)$		60 dB SNR $g(x, y)$		40 dB SNR $g(x, y)$	
	Q dB Seed 1	Q dB Seed 2	Q dB Seed 1	Q dB Seed 2	Q dB Seed 1	Q dB Seed 2
5×5	126.95	124.25	37.51	32.50	11.17	14.71
8×8	115.53	117.57	20.75	27.33	4.31	7.34
16×16	111.56	98.45	19.09	17.63	3.79	4.04

Table 5.6 Seed 1 and Seed 2 signify that all the random images used to produce the results in the respective columns were derived from the same random number generator seed.

To determine what effect increasing levels of contamination had on Q an evaluation of the ability of the algorithm to deconvolve images of varying SNRs has been conducted. In this evaluation four 31×31 pixel convolutions, consisting of 16×16 random images and psfs, and four 15×15 convolutions, comprising 8×8 random images and psfs, were corrupted with varying levels of contamination to give the resulting convolutions 60 dB, 50 dB, 40 dB, and 30 dB SNRs. In Fig. 5.9 the quality of the $\hat{f}(x, y)$ obtained from the four 15 × 15 pixel convolutions are plotted against the convolution SNRs. This plot clearly shows that as the

SNR of the convolution decreases the quality of the $\hat{f}(x,y)$ obtained from the consistent algorithm decreases. There are three factors which contribute to this decline in Q as the SNR decreases. The first of these is the dependence of the accuracy of the set of equations, defining $\hat{f}(x,y)$, on the accuracy of the zero samples used to produce the equations. As the SNR of the convolution decreases, the zero samples used to produce the set of equations become more inaccurate and as a result Q declines as the accuracy of the set of equations declines. The second factor is the decrease in the likelihood of obtaining a well-conditioned set of equations as the SNR decreases. The final factor is that the amount of zero sheet merging tends to increase as the SNR of the convolution decreases (see 4.4.3). The increased incidence of zero sheet merging makes it more difficult to correctly categorise the sections of zero sheet into their respective components and as a result the reconstruction Q declines.

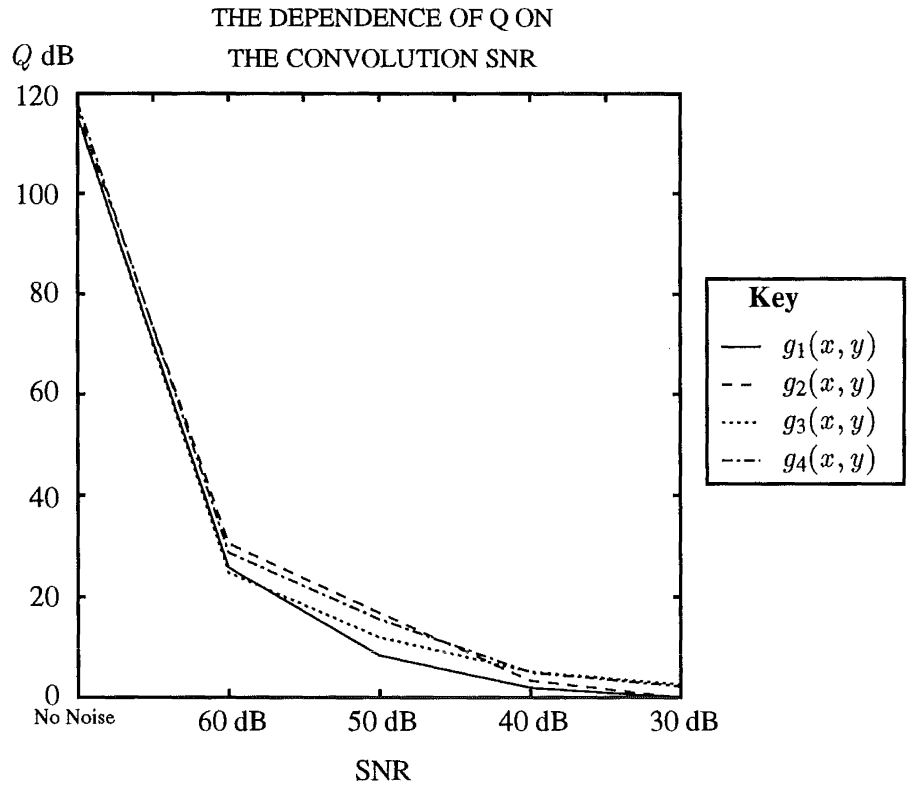

Figure 5.9 The dependence of the $\hat{f}(x,y)$ Q on the $g(x,y)$ SNR for four different $f(x,y)$

The final characteristic of the convolution investigated to determine what effect it had on $\hat{f}(x,y)$ Q was the size of $h(x,y)$. To determine how the size of $h(x,y)$ effects the $\hat{f}(x,y)$ Q, four 16×16 pixel images were convolved with 16×16, 8×8, and 4×4 random psfs to give a set of blurred images, with different size psfs, that could be deconvolved by the consistent algorithm. All of the convolutions described above were contaminated so that the resulting images had SNRs of 60 dB. The results from this investigation are presented in Table 5.7 and it can be seen that the $\hat{f}(x,y)$ Q declines as the psf size increases. The reason for this result is that the number of zero tracks comprising the convolutions zero sheet slice increases as the psf size increases and as a consequence the incidence of zero track merging also increases. As explained above, an increase in zero track merger decreases $\hat{f}(x,y)$ Q because the algorithm finds it more difficult to correctly classify all the zero tracks.

In summary, the results in this section indicate that the quality of the $\hat{f}(x,y)$ obtained from the consistent algorithm decreases as either the original image size, psf size, or contamination

THE DEPENDENCE OF Q ON PSF SIZE

psf size	Q dB $f_1(x,y)$	Q dB $f_2(x,y)$	Q dB $f_3(x,y)$	Q dB $f_1(x,y)$
4×4	30.65	32.46	36.54	30.85
8×8	22.38	21.03	22.32	15.20
16×16	9.94	12.67	7.93	9.14

Table 5.7 The subscripts denote that the convolutions were derived from four different original 16×16 pixel images. Each of the original images was convolved with the psf of the indicated size and the resulting convolution was contaminated to give it a 60 dB SNR.

level increase. The reasons for this decrease in Q are that an increase in the above parameters either makes the set of equations defining $\hat{f}(x,y)$ increasingly ill-conditioned, increases the inaccuracy of the set of equations and/or increases the likelihood of zero sheet merger. The results in this section appear to indicate that the consistent algorithm in its present form is unlikely to be able to obtain useful $\hat{f}(x,y)$ from convolutions comprising original images larger than 16×16 pixels when the convolution SNR is less than 40 dB. To pictorially demonstrate to the reader that the consistent algorithm can successfully deconvolve images of extent up to 16×16 pixels and with convolution SNRs of as low as 40dB the $\hat{f}(x,y)$ obtained from a blurred and contaminated image of the letter G and a blurred and contaminated random image are shown in Fig. 5.10. The image of the letter G has been included within this figure as its form is much simpler to recognise than that of a random image. The ability of the consistent algorithm to obtain useful $\hat{f}(x,y)$ can be seen clearly in both the reconstructions from the 60dB SNR convolutions (Fig. 5.10 (c) and (g)). With convolution SNRs of 40dB the consistent algorithm is able to produce just recognisable reconstructions.

5.5 COMPARISON OF ALGEBRAIC AND FOURIER RECONSTRUCTION

In §5.4 it was shown that the accuracy of the consistent algorithm decreases as the size of the original image in the convolution to be deconvolved increases. As explained in §5.4 this decrease in the quality of $\hat{f}(x,y)$ occurs because the number of zero samples required, by the algebraic reconstruction technique to form the set of equations defining $\hat{f}(x,y)$, increases with image size and it therefore becomes more difficult to obtain a well-conditioned set of equations. It is also shown in §5.6 that the algebraic reconstruction technique is computationally very expensive with the amount of computation required rising sharply with small increases in image size. These difficulties with the algebraic reconstruction technique prompted the author to investigate whether the Fourier based reconstruction technique, described in §4.5.1, could be used in conjunction with the recognition and separation scheme to quickly obtain good quality reconstructions from larger images. As indicated in the theoretical comparison of the two reconstruction techniques in §4.5.3, the computational requirements of the Fourier technique are significantly less than those of the algebraic reconstruction technique and as the Fourier technique does not involve solving a large system of linear equations it is possible that this technique may provide superior reconstructions. The results from this investigation and the modifications that had to be made to the consistent algorithm to accommodate the Fourier reconstruction technique are discussed in this section.

Before attempting to use the Fourier based reconstruction technique to obtain an $\hat{f}(x,y)$ from $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ it was necessary to make a few modifications to the recognition and separation algorithm which produces $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$. These modifications were required because the Fourier reconstruction technique normally requires that specific zeros are sampled from the zero sheet,

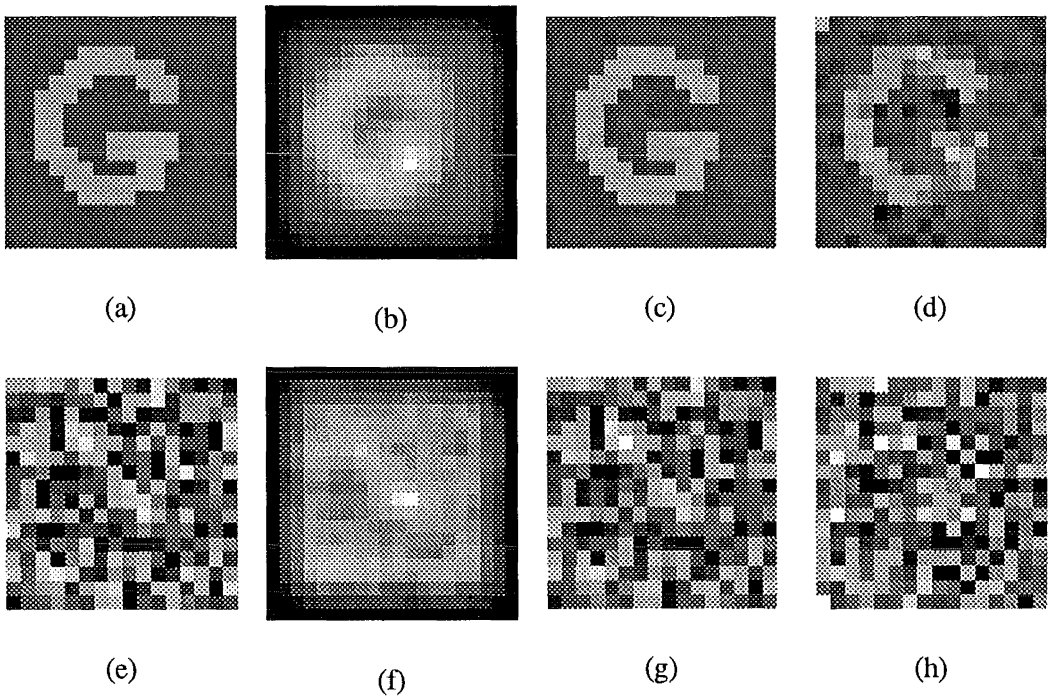

Figure 5.10 (a) a 16×16 pixel original image of the letter G, (b) the image in (a) after it has been blurred by the psf in Fig. 5.2 (c), (c) the reconstruction (Q = 25.10 dB) obtained from the consistent algorithm when (b) had a 60 dB SNR, and (d) the reconstruction (Q = 11.70 dB) obtained when (b) had a 40 dB SNR. (e) a 16×16 pixel random image, (f) the image in (e) after it has been blurred by the psf in Fig. 5.2 (c), (g) the reconstruction (Q = 19.08 dB) obtained from the consistent algorithm when (f) had a 60 dB SNR, and (h) the reconstruction (Q = 4.01 dB) obtained when (f) had a 40 dB SNR.

although as was described in §4.5.3 recent developments have given this technique a limited amount of freedom with regard to sampling the zero sheet (Satherley and Parker, 1993). As the Fourier reconstruction technique cannot use zero samples from anywhere on the zero sheet the recognition and separation scheme can no longer remove zeros from $Z\{\mathcal{G}(\zeta,\gamma)\}$ which it is unable to classify as belonging to the psf or image. The zeros normally removed may be those zeros required by the Fourier reconstruction technique. As a consequence, the separation routine was adapted so that it no longer discards zeros that belong to the zero tracks closest to the psf estimate's zero tracks if they are not heading in the same direction as the psf estimate's zero tracks (i.e. if $\theta_p^k < \pi/4$). The separation routine also no longer discards the zeros belonging to the sections of track that have a length less than 5 % of the total track length. The modified algorithm therefore splices together those sections of $\mathcal{G}(\zeta,\gamma)$ zero track that are closest to the psf estimate's zero tracks and which head in the same direction, however those sections of zero track which do not head in the same directions as the psf estimate's zero tracks are not touched by the separation routine.

The two 12×12 and two 20×20 pixel convolutions in the set of test images described in §5.3.1 were processed with the consistent algorithm incorporating the Fourier reconstruction technique to determine whether useful reconstructions could be obtained. In Table 5.8 the quality of the $\hat{f}(x,y)$ obtained in the situation where the convolutions had a 60 dB SNR or were uncontaminated are given. Also included within this table, for comparison purposes, are the results obtained from the same images when the consistent algorithm incorporating the algebraic reconstruction technique was used to perform the deconvolution. The results in this table show that in the uncontaminated situation the Fourier reconstruction technique can provide

$\hat{f}(x,y)$ with a Q comparable to that obtained with the algebraic reconstruction technique, i.e. the accuracy of the reconstructions were similar to the accuracy of the number representation scheme in the computer used. In the contaminated situation, however, the Fourier technique was unable to provide useful $\hat{f}(x,y)$ (i.e. $Q \leq 0$) from any of the convolutions, whereas the algebraic reconstruction technique provided $\hat{f}(x,y)$ of a very useful quality.

COMPARISON OF FOURIER
AND ALGEBRAIC RECONSTRUCTION TECHNIQUES

No Contamination				
Technique	Q dB 12×12 $g_1(x,y)$	Q dB 12×12 $g_2(x,y)$	Q dB 20×20 $g_1(x,y)$	Q dB 20×20 $g_2(x,y)$
Fourier	118.69	92.31	107.56	102.32
Algebraic	118.16	119.73	113.19	104.34
60 dB SNR $g(x,y)$				
Fourier	≤ 0	≤ 0	≤ 0	≤ 0
Algebraic	20.75	27.32	19.09	12.63

Table 5.8 Fourier = Fourier reconstruction technique. Algebraic = Algebraic reconstruction technique. The subscripts denote that the convolutions were derived from different original images.

The reason why the Fourier reconstruction technique was unable to provide useful $\hat{f}(x,y)$ in the contaminated situation was because the separation algorithm was unable to accurately and completely classify all the zeros comprising $Z\{\mathcal{G}(\zeta,\gamma)\}$ as belonging to the psf or image. The $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ slice produced by the separation algorithm therefore contained sections of zero track which belonged to $Z\{\mathcal{H}(\zeta,\gamma)\}$. When the Fourier technique sampled $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ to reconstruct $\hat{F}(u,v)$ the technique had no choice but to sample those zeros which actually belonged to the psf. As these zero samples of $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ that actually belong to $Z\{\mathcal{H}(\zeta,\gamma)\}$ could be positioned anywhere in Z space the reconstructed $\hat{F}(u,v)$ tended to be grossly inaccurate and as a result the $\hat{f}(x,y)$ obtained were of no use. The Fourier reconstruction technique is very sensitive to the accuracy of the zero samples as it only takes one grossly inaccurate zero sample to produce a $\hat{F}(u,v)$ and $\hat{f}(x,y)$ which is grossly inaccurate. The accuracy of the Fourier reconstruction technique may be able to be improved by using the new Fourier reconstruction technique that allows some flexibility over the position of the zero samples or by designing a new zero sheet separation and separation scheme, specifically for the Fourier reconstruction technique, that allows the zeros of $Z\{\mathcal{G}(\zeta,\gamma)\}$ to be classified more accurately. The likelihood of any new technique being able to completely avoid sampling incorrectly classified zeros or being able to correctly classify all the zeros of $Z\{\mathcal{G}(\zeta,\gamma)\}$ is however very small.

The consistent algorithm incorporating the algebraic reconstruction technique does not suffer from the above problem because the sampling flexibility of the technique allows the separation algorithm to remove those zeros from $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ which it cannot classify with a reasonable certainty. As a consequence, the zeros sampled from $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ by the algebraic reconstruction technique are less likely to be those that belong to $Z\{\mathcal{H}(\zeta,\gamma)\}$. In addition the algebraic reconstruction technique obtains $\hat{f}(x,y)$ by solving an overdetermined system of linear equation in a least square sense, hence a small number of inaccurate zero samples are unlikely to prevent a useful $\hat{f}(x,y)$ being obtained. This study has shown that the most suitable reconstruction technique for the consistent algorithm is the algebraic reconstruction technique as, although it is computationally much more expensive than the Fourier reconstruction technique, it is very robust with respect to the incorrect classification of the zero samples within $Z\{\mathcal{G}(\zeta,\gamma)\}$.

5.6 COMPUTATIONAL REQUIREMENTS

The computational requirements of the consistent algorithm have been experimentally determined to allow comparison with that of other standard deconvolution techniques. In addition, the computational cost of the various components comprising the algorithm was evaluated to determine which component contributed most to the computational cost, how the computational cost scaled with original image size, and whether the scaling with image size occurred at the same rate as that predicted by the theory. The three key components within the consistent algorithm whose computational cost had to be evaluated were the algorithm used to compute the images' zero sheets, the recognition and separation algorithm, and the algebraic reconstruction algorithm. The reconstruction component of the algorithm used a single value of r_γ. The results of others predict that for large images (i.e. $> 16\times16$) the computation time required to compute the zeros of a $N \times N$ image should be $O((N-1)^2)$ (Jenkins and Traub, 1972) and the reconstruction of the image from its zeros should be $O((N)^6)$ (Dongarra *et al.*, 1979; Izraelevitz and Lim, 1987). The author predicts that the computational requirements of the recognition and splicing algorithm should be $O(N)$ in the situation where the extents of the psf are much less than those of the original image, since every psf zero track is compared to every convolution zero track. It is also worth noting that Izraelevitz and Lim (1987) report for their algorithm, which solves the phase retrieval problem by using the Z transform zeros and algebraic reconstruction, that for images up to 16×16 pixels the zero computation is computationally more expensive than the algebraic reconstruction and for larger images the situation reverses.

To evaluate the computational cost of the consistent algorithm the set of uncontaminated and contaminated 9×9, 12×12, and 20×20 pixel convolutions described in §5.3.1 (original image sizes of 5×5, 8×8, and 16×16) were deconvolved using the algorithm and the time to complete the computation for each component was recorded. The computer used to determine the computational requirements of the algorithm was a VAX station 3100. All of timing results have been normalised with respect to the time required to compute an 8×8 FFT on this machine so that interested readers can derive the computational requirements of this algorithm on other machines. In Table 5.9 the observed computation times for the different sections of the algorithm are tabulated. The computation times given in Table 5.9 for each size of image are the mean values of all the computation times observed for all the images of that particular size. The results in this table demonstrate that deconvolving a blurred image by using the consistent algorithm is computationally very expensive. To deconvolve the smallest convolutions, which consisted of 5×5 pixel original images, required the same amount of computation as that required to compute 4943 8×8 FFTs. If the computational cost of the consistent algorithm is compared to that of the Wiener filter the large computational cost of the former technique becomes even more apparent. Consider the case where the Wiener filter is used to deconvolve a blurred image consisting of a 8×8 original image and psf of much smaller extent. The time taken for the Wiener filter to obtain a reconstruction in this situation is equivalent to the time taken to compute approximately three 8×8 FFTs, i.e. one to compute $G(u,v)$, one to compute $H(u,v)$, and one to compute $\mathsf{F}^{-1}\{\hat{F}(u,v)\}$. On comparing the Wiener filters computational cost of approximately three 8×8 FFTs to the consistent algorithm's 8593 8×8 FFTs it can be seen that the latter is very expensive. It is necessary to note, however, that the Wiener filter is often applied many times to establish the filter constant that produces the best reconstruction.

As explained above for the purposes of this comparison r_γ was held constant and as a consequence only one reconstruction was obtained by the algebraic decomposition algorithm. In the situation where r_γ is incremented from 1.0 to 1.5 in steps of 0.025 and $\hat{f}(x,y)$ are

COMPUTATIONAL REQUIREMENTS OF THE CONSISTENT ALGORITHM

$f(x,y)$ size	Zeros			Separation			Reconstruction			Total	
	FFTs	s	%	FFTs	s	%	FFTs	s	%	FFTs	s
5×5	4371	88	88	497	10	10	75	2	2	4943	99
8×8	6855	138	80	745	15	8	993	20	12	8593	173
16×16	17933	361	25	2483	50	4	49975	1006	71	70391	1417

Table 5.9 Key: Zeros = Zeros Computation, Separation = Zero sheet recognition and separation, Reconstruction = Algebraic Reconstruction, FFT = computation time in 8×8 FFTs, s = computation time in seconds on VAX station 3100, and % = percentage of total computation time.

obtained for each increment in r_γ, in an attempt to find the best $\hat{f}(x,y)$ (see §5.3.1), the consistent algorithms computational cost is equivalent to that of computing 1×10^6 8×8 FFTs (5.7 hrs) in the case of a 16×16 pixel original image. The enormous computational expense of this reconstruction method prompted the investigation in §5.3.1 which attempted to determine whether good quality image estimates could be reliably obtained by incrementing r_γ in larger steps.

As the image size increases it can be seen in Table 5.9 that the computational cost of the consistent algorithm rises dramatically. As the images used here are small (i.e. $\leq$ 16×16), in comparison to those used by others to derive the expressions which describe how the computation required by different sections of the algorithm scales with image size, caution must be exercised when comparing the experimental results presented here to those of others. On computing the scaling factor for the zero computation algorithm from the experimental results in Table 5.9 it was found that the algorithm's computational cost was scaling approximately $O(N)$ which is different to the $O((N-1)^2)$ reported by Jenkins and Traub (1972). The difference in the above results is most likely due to the small size of the images used here. The experimental results in Table 5.9 also indicate that the computational expense of the algebraic reconstruction technique, for the small images used here, was scaling somewhere between $O(N)^5$ and $O(N)^6$ which agrees reasonably well with the previously reported result of $O(N)^6$ (Dongarra *et al.*, 1979; Izraelevitz and Lim, 1987). In addition the results demonstrate, as the scaling factors above predict, that the algebraic reconstruction technique dominates the computational requirements of the algorithm as the image size increases, while the zero computation algorithm becomes less demanding in comparison, i.e. for the 5×5 image the percentage of the total computation time that could be attributed to the zero computation, recognition and separation, and reconstruction algorithms was 88 %, 10 % and 2 % respectively, while for the 16×16 image the respective proportions were 25 %, 4 %, and 71 %. This result, describing how the computational cost of the different components within the algorithm scale with image size, confirms the finding of Izraelevitz and Lim (1987) that the algebraic reconstruction algorithm becomes more demanding than the zero computation as the image size increases.

In summary, the above results have shown that the consistent algorithm is computationally a very expensive algorithm. The computational cost of this algorithm vastly exceeds that required by the Wiener filter and as the image size increases the computational cost rises dramatically because the cost of the reconstruction component of the algorithm scales $O(N)^6$. The results above indicate that the consistent algorithm must either provide reconstructions with a superior Q to those of other techniques, or which are less sensitive to the $\hat{h}(x,y)$ accuracy, if the excessive computational cost of using the consistent algorithm is to be justified.

5.7 COMPARISON TO THE WIENER FILTER

In §5.1 it was explained that one of the motives for developing the consistent algorithm was that it was believed that the algorithm would provide solutions to the standard deconvolution problem which were less sensitive to errors in $\hat{h}(x,y)$ than the existing deconvolution techniques. To determine whether this aim has been achieved a study comparing the accuracy of the $\hat{f}(x,y)$ obtained from both the Wiener filter and consistent algorithm, in the situation where $\hat{h}(x,y) \neq h(x,y)$, has been conducted. To simulate the errors in $\hat{h}(x,y)$ in this study, random contamination was added to $h(x,y)$ in a manner which gave the resulting $\hat{h}(x,y)$ one of the SNRs required by the study. The effect of this contamination on the zero sheet of $h(x,y)$ is to perturb the zeros, as described in §4.4.1, and the consistent algorithm must therefore match these perturbed zeros to those in $Z\{\mathcal{G}(\zeta,\gamma)\}$ that correspond to $Z\{\mathcal{H}(\zeta,\gamma)\}$. Another form of psf error that could have been considered, and which is likely to arise in practice, is the error that arises when the form of the psf is known but an accurate estimate of the psf extent is unknown i.e. it is known that $g(x,y)$ is out of focus but the diameter of the disc which describes the blurring is unknown. An inaccurate estimate of the psf extent has the effect of perturbing the Fourier transform zeros of the psf in a manner described by the scaling property in Table 2.1 and as a consequence the Z transform zeros are also perturbed in a similar manner. As the psfs dealt with here are very small (i.e. $\leq 16\times16$) and the effect of both types of psf estimate error is simply to perturb the zeros of $Z\{\mathcal{H}(\zeta,\gamma)\}$ only the case of random psf error was considered in this study.

In this study the set of blurred images described in §5.3.1 were deconvolved by both the Wiener filter and the consistent algorithm. The $\hat{h}(x,y)$ given to both deconvolution techniques were versions of $h(x,y)$ that had random contamination added to them to give the resulting psf either a 60 dB, 50 dB, 40 dB, 30 dB, or 20 dB SNR. In this study the Wiener filter constant Φ (see §3.2.2) was initially set to the noise-to-signal ratio of the convolution, Φ was then adjusted until the $\hat{f}(x,y)$ with the largest possible Q was obtained. In Table 5.10 representative results from this study for 60 dB and 40 dB SNR 9×9 pixel contaminated convolutions are tabulated alongside those from 60 dB and 40 dB SNR 20×20 pixel contaminated convolutions. The results in Table 5.10 for the smaller convolution with a 60 dB SNR show that the consistent algorithm is able to provide $\hat{f}(x,y)$ with a Q superior to those obtained by Wiener filtering. It can also be seen that the $\hat{f}(x,y)$ Q remains reasonably stable with the decrease in SNR of $\hat{h}(x,y)$ while the Q of the $\hat{f}(x,y)$ obtained from the Wiener filter steadily declines as the SNR of $\hat{h}(x,y)$ decreases. This first result therefore demonstrates that the consistent algorithm has met its design goal as it is less sensitive to psf error than the Wiener filter.

When either the SNR of $g(x,y)$ decreases or the size of the original image becomes large the results in Table 5.10 demonstrate that the Wiener filter is always able to provide a $\hat{f}(x,y)$ with a Q superior to that provided by the consistent algorithm. The results for the 60 dB SNR 20×20 pixel convolution indicate that the consistent algorithm is still less sensitive to the psf accuracy than the Wiener filter. The Wiener filter is able to provide $\hat{f}(x,y)$ with such a high Q, however, that even the Wiener filter's decline in $\hat{f}(x,y)$ Q with decreasing $\hat{h}(x,y)$ SNR is not sufficient to let the consistent algorithm provide superior $\hat{f}(x,y)$ when the SNR of $\hat{h}(x,y)$ is 20 dB, i.e. for $\hat{h}(x,y)$ SNRs of 60 dB and 40 dB the consistent algorithm's $\hat{f}(x,y)$ Q were 7.45 dB and 6.42 dB respectively, while the Wiener filter's were 31.60 dB and 12.00 dB respectively. The set of images used to derive the above results were all blurred by the same psf, however results are presented in §5.8, for four random 8×8 pixel images convolved with four different 8×8 pixel psfs, which demonstrate that the consistent algorithm's robustness with respect to the psf error is not dependent on the form of the psf.

COMPARISON OF THE CONSISTENT ALGORITHM
AND WIENER FILTER

9×9 convolution						
$\hat{h}(x,y)$ SNR	60 dB SNR $g(x,y)$			40 dB SNR $g(x,y)$		
	Zeros Q dB	Wiener Q dB	ΔQ	Zeros Q dB	Wiener Q dB	ΔQ
60	36.84	31.50	5.84	13.02	- 18.60	-5.58
50	36.79	28.10	8.69	12.59	18.50	-5.91
40	36.82	24.00	12.82	12.37	19.50	-7.13
30	31.33	18.50	12.83	11.28	16.80	-5.52
20	31.48	14.10	17.38	8.06	13.80	-5.74
20×20 convolution						
$\hat{h}(x,y)$ SNR	60 dB SNR $g(x,y)$			40 dB SNR $g(x,y)$		
	Zeros Q dB	Wiener Q dB	ΔQ	Zeros Q dB	Wiener Q dB	ΔQ
60	7.45	31.60	-24.15	2.42	16.80	-14.38
50	7.10	30.30	-23.20	2.60	17.00	-14.40
40	7.80	24.40	-16.60	3.62	16.30	-12.68
30	7.55	19.10	-11.55	3.39	15.90	-12.51
20	6.42	12.00	-5.58	0.48	12.10	-11.62

Table 5.10 ΔQ = Zeros Q dB - Wiener Q dB.

In summary, the above results demonstrate that the consistent algorithm is robust with respect to the error on $\hat{h}(x,y)$ and thus the algorithm does meet its design goals. The quality of the $\hat{f}(x,y)$ are, however, only superior to those obtained by the Wiener filter in the situation where the SNR of the convolution is high and the original image is of a small extent. For higher SNRs or images of a large extent the consistent algorithm finds it difficult to obtain a well-conditioned sets of equations (see §5.4) and Q drops to a level which is significantly lower than that which can be obtained by Wiener filtering. As noted in §3.2.2 the Wiener filter is not considered to be an especially good deconvolution technique and as a consequence these results demonstrate that the consistent algorithm cannot be considered, at present, to be an especially good deconvolution technique for images which have extent larger than 5×5 or for convolutions with SNRs of lower than 60 dB.

5.8 COMPARISON TO POINT ZERO MATCHING ALGORITHM

A new zero-based algorithm which should also be able to provide a consistent solution to the standard deconvolution problem has recently been reported by Parker (1992) (Bones *et al.*, 1993). This zero based algorithm, known hereafter as the zero matching algorithm, is an adapted version of the algorithm originally developed by Parker and Bones (1992) to aid the convergence of the iterative phase retrieval algorithms described in §3.5.2. The zero matching phase retrieval algorithm is described fully in §6.1.1 and so only a brief review of the zero matching algorithm is given here.

The zero matching algorithm uses the same zero sheet separation principle which the consistent algorithm uses to obtain a consistent solution to the deconvolution problem. The zero matching algorithm first computes the zeros of $\hat{\mathcal{H}}(\zeta,\gamma)$ and $\mathcal{G}(\zeta,\gamma)$ which respectively define the Fourier spectra of the psf and convolution along lines of constant u or v in the real Fourier plane which characterise the DFT sampling grid (see §4.5.1). As $G(u,v)$ is

approximately the product of $F(u,v)$ and $H(u,v)$ the zeros that best approximate those of $F(u,v)$ along these lines of constant u or v can be found by invoking the consistency condition, i.e. for a particular u_n those point zeros of $G(u_n,v)$ closest to the zeros of $\hat{H}(u_n,v)$ are the best estimate of the zeros of $H(u_n,v)$ and the remaining zeros are the best estimate of the zeros characterising $F(u_n,v)$. It is therefore possible to obtain the set of zeros which define $\hat{F}(u,v)$ on the DFT sampling grid and obtain $\hat{f}(x,y)$ by using the IDFT. The major difference between the zero matching algorithm and the consistent algorithm are that the zero matching algorithm only attempts to match point zeros while the consistent algorithm attempts to match sections of zero track, and the zero matching uses the Fourier reconstruction technique while the consistent algorithm uses the algebraic reconstruction technique. Preliminary results presented by Parker (1992) suggest that the zero matching algorithm may be able to deconvolve larger images than the consistent algorithm due to the fact that it uses the Fourier reconstruction technique. The author, however, expected that the zero matching algorithm may struggle to obtain a useful $\hat{f}(x,y)$ in the situation where the convolution is badly contaminated or the psf estimate is poor. To determine if the above beliefs were correct a study comparing the performance of the two algorithms was conducted.

A set of blurred images of varying sizes and levels of contamination were deconvolved by the algorithms with the aid of psfs of varying accuracies. The images used in this study were four 31×31 and four 15×15 pixel convolutions formed from 16×16 and 8×8 pixel random images and random psfs respectively. The convolutions were corrupted with contamination to give them SNRs of 60 dB, 50 dB, 40 dB, 30 dB, or 20 dB and the psfs used to obtain the $\hat{f}(x,y)$ had SNRs of either ∞, 60 dB, or 40 dB. In the following results, the Q values quoted are the mean of the four Q values obtained from the four different blurred images of each size.

COMPARISON OF THE CONSISTENT ALGORITHM
AND ZERO MATCHING ALGORITHM
WHEN THE PSF IS ACCURATE

	15×15 convolution			31×31 convolution		
$g(x,y)$ SNR	Consistent Q dB	Zero Match Q dB	ΔQ dB	Consistent Q dB	Zero Match Q dB	ΔQ dB
∞	116.45	110.79	5.66	104.90	100.40	4.50
60	27.37	31.00	-3.63	9.92	22.86	-12.94
50	13.10	21.59	-8.49	1.25	13.03	-11.78
40	3.83	14.34	-10.51	0.61	6.29	-5.86
30	1.18	6.58	-5.40	0.20	0.32	-0.12

Table 5.11 Consistent = Consistent Algorithm, Zero Match = Zero Matching Algorithm, and ΔQ = Consistent Q dB - Zero Match Q dB.

In the situation where the actual psf was used by the algorithms to deconvolve the blurred images, with the SNRs stated above, the results tabulated in Table 5.11 were obtained. These results clearly show that, when the convolution is contaminated, the zero matching algorithm is always able to provide image estimates of a superior quality to those provided by the consistent algorithm. On comparing the computational cost of the two algorithms it was found that the computational cost of the zero matching algorithm was significantly less than that of the consistent algorithm. For the 15×15 pixel convolution the normalised computational cost of the zeromatching algorithm was equivalent to computing 335 8×8 FFTs, while the cost of the consistent algorithm was equivalent to 8593 8×8 FFTs. For the 31×31 pixel convolution the computational costs were 4272 8×8 FFTs and 70391 8×8 FFTs for the zero matching and

consistent algorithms respectively. When the psf was inaccurate the results presented in Table 5.11 were obtained from the 15×15 pixel convolution by the algorithms. These results show that as the psf becomes inaccurate both algorithms are able to provide image estimates of a quality which is largely unaffected by the decrease in accuracy of the psf. The results also show that the reconstructions provided by the zero matching algorithm are less sensitive to the psf estimate inaccuracies than the consistent algorithm.

COMPARISON OF THE CONSISTENT ALGORITHM
AND ZERO MATCHING ALGORITHM
WHEN THE PSF IS INACCURATE

	Consistent Algorithm			Zero Matching Algorithm		
$g(x,y)$ SNR	$\hat{h}(x,y)$ SNR ∞	$\hat{h}(x,y)$ SNR 60 dB	$\hat{h}(x,y)$ SNR 40 dB	$\hat{h}(x,y)$ SNR ∞	$\hat{h}(x,y)$ SNR 60 dB	$\hat{h}(x,y)$ SNR 40 dB
∞	116.45	116.44	116.33	110.79	110.79	110.79
60	27.37	27.76	27.72	31.00	31.00	31.12
50	13.10	12.77	14.18	21.59	21.59	21.82
40	3.83	3.69	3.49	14.34	14.34	14.60
30	1.18	0.83	1.15	6.58	6.35	7.62

Table 5.12

The above results demonstrate that the author's belief that the consistent algorithm would be able to provide image estimates of a quality superior to those provided by the zero matching algorithm, when the convolution SNR is low or the psf is inaccurate, was incorrect. It appears that, for the convolution and psf estimate SNRs investigated here, the zero matching algorithm is able to obtain a set of image estimate zeros of an accuracy which allows the Fourier reconstruction method to be used to obtain a good quality image estimate. At convolution SNRs of 60 dB and 50 dB the consistent algorithm is able to obtain accurate estimates of the zero tracks of the image (see Fig. 5.4) and so the algebraic reconstruction technique must be the section of this algorithm which is responsible for its inferior performance. Further evidence in support of this assertion is the very small change in reconstruction quality with increasing psf estimate inaccuracy (see Table 5.12), which suggests that the consistent algorithm is able to recognise the zero tracks of the image with a good degree of accuracy. It appears, therefore, that the ill-conditioned nature of the set of equations defining $\hat{f}(x,y)$ is responsible for the inferior reconstructions provided by consistent algorithm. As shown in §5.4 these equations become increasingly ill-conditioned as the image size increases or as the convolution SNR decreases.

5.9 ANTICIPATED DIFFICULTIES WITH PRACTICAL APPLICATION

There are two properties of practical convolutions which the author anticipates may adversely effect the ability of the consistent algorithm to obtain useful $\hat{f}(x,y)$ from practical $g(x,y)$. The first property is that a large number of practical images are in effect not of finite extent and as a consequence the blurred image is in practice usually a truncated version of the actual convolution. The second property is that the characteristics of most practical imaging systems result in appreciable levels of contamination being added to the convolution. The implications of these properties of practical convolutions are now discussed in this section and where possible methods that could be used to overcome the difficulties that these properties pose are

discussed.

The deconvolution techniques described in §3.2 and the consistent algorithm all assume that it is possible to record the complete convolution $g(x,y)$, however in practice it is usually impossible to measure $g(x,y)$ and only a truncated version of the convolution, $g_t(x,y)$, is available. This truncation arises when the size of the original image is in effect infinite, i.e. the extent of the image is much greater than the extent of the recording device. In this situation the pixels in the recorded image $g_t(x,y)$ contain a contribution from the blurred pixels of the original image which are outside the extents of $g_t(x,y)$. This contribution to $g_t(x,y)$ from neighbouring pixels can induce rather serious artifacts in the $\hat{f}(x,y)$ obtained and these artifacts tend to be particularly bad when the neighbouring pixels introduce rather sharp changes in the pixel intensity near the perimeter of $g_t(x,y)$ (Bates and McDonnell, 1986). It is likely that this form of image truncation would also introduce artifacts in the $\hat{f}(x,y)$ obtained by the consistent algorithm as the distortion of the high frequency spectrum of the convolution which produces these artifacts is also likely to perturb the zero sheets of the component images. The fact that the $Z\{\mathcal{G}_t(\zeta,\gamma)\}$ slice being used to obtain $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$ has less zero tracks than $Z\{\mathcal{G}(\zeta,\gamma)\}$ should not be a problem, as the zero tracks pertaining to $Z\{\mathcal{H}(\zeta,\gamma)\}$ should still be present and the remaining zero tracks should characterise a truncated version of $\hat{f}(x,y)$.

To reduce the effects of these artifacts it is desirable to gradually taper the magnitude of the pixels around the perimeter of the convolution to be convolved to zero, as this removes any sharp changes in the intensity near the perimeter (or equivalently reduces the extent of the spectral leakage in $G_t(u,v)$). There are two methods which can be used to give the convolution the required form: the first technique involves multiplying $g_t(x,y)$ by a window function, which tapers to zero at its perimeter (Bates and McDonnell, 1986; Brigham, 1988). The second method is to use the edge extension techniques of McDonnell and Bates (1975) (Bates and McDonnell, 1986) which attaches additional pixels to the perimeter of $g_t(x,y)$ in a manner that results in a new convolution, $\breve{g}(x,y)$, that tapers off to zero at its perimeter. Bates and McDonnell (1986) consider the edge extension methods to be more sophisticated than windowing as it does not alter any of the recorded information, i.e. it does not alter $g_t(x,y)$. Both of these techniques could most likely be used with the consistent algorithm if it was found that the $\hat{f}(x,y)$ obtained from $g_t(x,y)$ possessed the artifacts normally associated with truncation.

If the consistent algorithm is to obtain a solution to a deconvolution problem it is necessary for the level of contamination on the convolution to be minimal otherwise Eqn.(4.19) is not valid and as is shown in §5.4 the consistent algorithm fails to separate the components zero sheets, i.e. the convolution SNR must be greater than about 30 dB. In practice most practical images are going to be corrupted by an appreciable amount of contamination, unless of course the image was obtained from a very high quality imaging instrument, i.e. not a standard still or video camera. The reason why it is likely that the contamination would be appreciable is that the lens systems in most practical imaging instruments introduce geometric distortions (Bates and McDonnell, 1986; Lee, 1990) and have a finite depth of field (see Fig. 2.8), which cause non-isoplanatic blurring. As explained in §2.3.2 these departures from isoplanatic blurring are included within the contamination term within the expression that models the blurred image (i.e. Eqn.(2.49)). In addition, as discussed in §2.3.2, it is likely that most practical blurred images have been corrupted by some of the following forms of contamination: nonlinearities associated with the recording medium or the scanning process used to digitise the image, recording noise, quantisation noise, film grain noise, truncation noise associated with convolution truncation, and physical defects such as dust and scratches. All of the above forms of contamination must also be included in the contamination term in Eqn.(2.49) and this term in practice is

therefore likely to be sizeable. Bates and McDonnell (1986) note than when the bandwidth of $f(x,y)$ is known it may be possible to reduce the effect of this contamination by filtering $g(x,y)$ in an manner that attenuates the contamination frequency components that are above the bandwidth of $f(x,y)$, however the remaining contamination is still likely to be of a significant level. Thus, in most practical imaging situations it is questionable whether the contamination term would be small enough to allow the approximation to be made that the zero sheet of the contaminated convolution is approximately equal to the union of $Z\{\mathcal{F}(\zeta,\gamma)\}$ and $Z\{\mathcal{H}(\zeta,\gamma)\}$, i.e. Eqn.(4.19) could be invalid. As a consequence the consistent algorithm is unlikely to be able to solve the practical deconvolution problem unless the blurred image is of a higher quality than commonly encountered.

Another factor, linked to the effects of practical contamination, which is also likely to limit the application of the consistent algorithm is the fact that the algorithm, at present, can only deconvolve very small images. As shown in §5.4 the algorithm has trouble obtaining the linearly independent set of equations that define $\hat{f}(x,y)$ in the situation where the extent of $f(x,y)$ is greater than 16×16 or the level of contamination on $g(x,y)$ becomes too great. Thus, the consistent algorithm in its present form could only be used in practice in the situation where very small high quality $g(x,y)$ can be obtained. One of the few practical fields of imaging science where very small images are regularly handled is the area of astronomical imaging. It is also worth noting that when the images are small and have a high SNR the consistent algorithm, which inherently satisfies the consistency condition, is able to produce very useful $\hat{f}(x,y)$ from inaccurate psf estimates. In fact the results in §5.7 demonstrated that in some situations the consistent algorithm could obtain $\hat{f}(x,y)$ with a Q superior to those obtained by other standard deconvolution techniques such as Wiener filtering.

CHAPTER 6

ZERO-ASSISTED PHASE RETRIEVAL

As explained in §3.1 phase retrieval can be considered to be a special case of blind deconvolution. It is therefore possible, as was shown in §4.3, to use the principle that a convolution's zero sheet is the union of its components' zero sheets to obtain a zero-based solution to the uncontaminated phase retrieval problem. When the power spectrum or convolution becomes contaminated, however, the merging of these component zero sheets makes it impossible to use the simple zero-based deconvolution and phase retrieval techniques described in §4.3. It was suggested in §3.5.2.4 that it may still be feasible to use the zero-based deconvolution principle to retrieve the phase, if the zero-based techniques are used in conjunction with the iterative techniques described in §3.5.2. Iteration could be used to estimate the image or its conjugate reflection and then a zero-based standard deconvolution technique, such as the consistent algorithm, used to improve the estimate. In essence iteration turns what was blind deconvolution into standard deconvolution. In this section a new phase retrieval algorithm combining the consistent algorithm with the iterative phase retrieval algorithms is described. The performance of this new algorithm is discussed and compared to that of the similar zero-based algorithm recently developed by Parker and Bones (1992).

In the first section of this chapter the zero-based algorithms which have been developed to assist the iterative phase retrieval algorithms are described. The technique developed by Parker and Bones (1992) is first described in §6.1.1 and the author's technique is then described in §6.1.2. Following these descriptions of the zero-based algorithms, the performance of the author's new algorithm is described in §6.2.1 and some experimental results are presented. An indication of the computational expense of the author's algorithm is given in §6.2.2 in comparison to that of the HIO algorithm. Finally, in §6.2.3 the performance of Parker *et al*'s algorithm is compared to that of the author's algorithm to determine which algorithm offers superior performance in different situations.

6.1 ZERO-BASED ALGORITHMS TO ASSIST PHASE RETRIEVAL

6.1.1 Point Zero Matching Algorithm

To assist the iterative phase retrieval algorithms of Fienup in obtaining a solution to the phase retrieval problem, Parker *et al* have recently developed a zero-based technique which improves the iterative algorithm's convergence. The improvements in convergence observed with Parker *et al*'s algorithm were noted for contaminated power spectra with SNRs $\geq$ 40 dB and with original image extents up to 32 $\times$ 32 pixels (Parker and Bones, 1992). It was also noted by Parker *et al* that the zero-based algorithm had the beneficial property that it could be used to

break the iterative algorithms out of stagnation. In this section the zero-based phase retrieval algorithm developed by Parker *et al* is described as its performance is compared, in a later section, to that of the new phase retrieval assistance algorithm developed by the author. For a more complete and detailed description of Parker *et al*'s algorithm, the reader is referred to Parker (1994).

To use Parker *et al*'s zero-based phase retrieval assistance algorithm, the phase retrieval problem must first be converted from a blind deconvolution problem to a standard deconvolution problem. To accomplish this conversion, one of the iterative phase retrieval algorithms is first used to provide an estimate of either the image or its conjugate reflection. This initial image estimate, $\tilde{f}(x,y)$, allows the phase retrieval problem to be treated as a standard deconvolution problem with $\tilde{f}(x,y)$ assuming the role of $\hat{h}(x,y)$. To ensure that Parker *et al*'s algorithm can successfully improve the quality of $\tilde{f}(x,y)$, it is necessary for $\tilde{f}(x,y)$ to be of sufficient accuracy. A strategy for automatically determining when enough iterations have been performed has yet to be determined. It is anticipated by Parker (1994) and the author, however, that in practice the violate error ϵ_v (see Eqn.(3.27)) could be used to indicate when the image estimate has reached the required accuracy. The iterative algorithm chosen by Parker *et al* to provide $\tilde{f}(x,y)$ is the HIO algorithm of Fienup (1982), as this algorithm's performance is generally superior to that of other algorithms (see §3.5.2.3). Once $\tilde{f}(x,y)$ has been obtained from the HIO algorithm some further preprocessing is required before Parker *et al*'s algorithm can be applied. It is first necessary to use prior knowledge of the extents of $f(x,y)$ to truncate $\tilde{f}(x,y)$ and the inverse Fourier transform of the contaminated power spectrum (contaminated auto-correlation), denoted hereafter by the function $i(x,y)$, to the correct extent. The difference in the extents of $\tilde{f}(x,y)$ and $f(x,y)$ arises because of the contamination, and truncation is required to ensure that the image estimate's zero sheet has the correct number of zeros at each value of ζ. The above also applies for the contaminated auto-correlation. After the above truncation has been performed, $\tilde{f}(x,y)$ and the contaminated auto-correlation are in a form for which Parker *et al*'s algorithm can be applied and an improved image estimate can thereby be obtained.

The point zero-matching algorithm obtains an improved solution to the phase retrieval problem by using the principle that $Z\{\tilde{\mathcal{F}}(\zeta,\gamma)\}$ can be used, in conjunction with the consistency condition, to separate the zero sheet of the contaminated auto-correlation into the two separate zero sheets which characterise $f(x,y)$ and its conjugate reflection. In actual fact the point zero-matching algorithm does not attempt to separate the entire zero sheet representing the contaminated auto-correlation into its respective components, but attempts to separate the subset of zeros which characterise the Fourier transform of the auto-correlation on the real (u,v) plane. On the real (u,v) plane, the spectrum of the auto-correlation along a line of constant u or v is described by a 1-D polynomial. This 1-D polynomial is characterised by a set of point zeros which, due to the relationship

$$I(u_c,v) = F(u_c,v)F^*(u_c,v),$$

where u_c denotes the chosen value of u, can be divided into two subsets of point zeros which individually characterise the 1-D polynomials that describe the spectra of the image and its conjugate along the same line. This set of zeros that characterise $I(u,v)$ along the line are a subset of the infinity of zeros which represent the auto-correlation's zero sheet. The lines of constant u or v which the point zero-matching algorithm uses to separate the auto-correlation's zeros into those defining its component images are those lines which define the DFT sampling grid in Fourier space, i.e. the lines in Fourier space whose intersections define where an image's spectrum must be sampled if the image is to be faithfully reproduced (see §2.5.4). On

separating the above sets of zeros and obtaining those that characterise $\hat{f}(x,y)$, it is possible to obtain $\hat{f}(x,y)$ directly using the DFT (see §4.5.1).

To produce and separate the sets of zeros required to obtain $\hat{f}(x,y)$ the zero-matching algorithm first sets $\zeta = exp^{i2\pi n/2N}$, where $n = 0,2,4,\cdots,2N-2$ and $f(x,y)$ is a $N \times N$ image, in $\tilde{\mathcal{F}}(\zeta,\gamma)$ and $\mathcal{I}(\zeta,\gamma)$. The point zeros of the resulting 1-D polynomials $\tilde{\mathcal{F}}(\zeta_n,\gamma)$ and $\mathcal{I}(\zeta_n,\gamma)$, which correspond to those along lines of constant u, are then computed using CPOLY (Jenkins and Traub, 1972). As described in §4.3.3 the point zeros of $\mathcal{I}(\zeta_n,\gamma)$, should be symmetrical about the unit circle in the γ plane, however Parker and Bones (1992) note that some pairs of zeros situated on the unit circle are not superimposed as they should be but are separated around the unit circle. This artifact arises because the preprocessing used to truncate $i(x,y)$ to the correct extent has the adverse effect of making the contaminated power spectrum go negative. As a negative power spectrum is not allowed by definition, and may prevent correct separation of the zeros, Parker *et al* replace the pair of zeros by a double zero situated on the unit circle midway between where the pair of zeros used to be situated. Replacing the zero pair by a double zero makes $\mathcal{I}(\zeta_n,\gamma)$ completely positive and preserves the correct extent of $i(x,y)$ in image space. After correcting the point zero symmetries, the point zero-matching algorithm then applies the consistency condition and selects the point zeros of $\mathcal{I}(\zeta_n,\gamma)$ closest to those of $\tilde{\mathcal{F}}(\zeta_n,\gamma)$ as the best estimate of those pertaining to $\mathcal{F}(\zeta_n,\gamma)$. These selected point zeros can be used to reconstruct the 1-D polynomial that describes $\hat{F}(u_n,v)$. As described in §4.5.1 all of the 1-D polynomials characterising the $\hat{F}(\zeta_n,\gamma)$ are scaled by a different complex constant. The point zero-matching algorithm therefore produces a normalising polynomial from a line of constant v by setting $\gamma = exp^{i2\pi m/2N}$, where m is an integer in the range 0 to $2N-2$, and performing the point zero-matching process for the zeros of $\mathcal{I}(\zeta,\gamma_m)$ and $\tilde{\mathcal{F}}(\zeta,\gamma_m)$. Once the normalising polynomial has been determined, all the 1-D polynomials for constant u are scaled and the Fourier reconstruction technique is used to obtain $\hat{f}(x,y)$. To refine $\hat{f}(x,y)$, Parker and Bones (1992) then make this image the initial estimate for the HIO algorithm and perform further HIO iterations. Parker and Bones (1992) note that the best strategy for refining $\hat{f}(x,y)$ is to perform a few iterations of HIO and then apply ER iterations (see §6.2.1) until the image estimate has the required accuracy. The application of the ER algorithm has been found to prevent the error drifting upwards during the image estimate refinement process.

Using the point zero-matching algorithm to assist the convergence of the iterative phase retrieval algorithm is computationally inexpensive. For an $N \times N$ image the point zero-matching algorithm must find the zeros of $2N+2$ one-dimensional polynomials which involves $2N+2$ applications of CPOLY. The point zeros must then be matched and the 1-D polynomials that characterise $\hat{F}(u,v)$ must be determined and sampled appropriately. Both of these functions require very little computation. To reconstruct $\hat{f}(x,y)$ requires only a single application of an $N \times N$ IFFT. Parker and Bones (1992) note that one application of point zero-matching is equivalent to approximately forty iterations of HIO for original images of extent 32×32 pixels.

6.1.2 Zero Track Matching Algorithm

The author's consistent algorithm, which was originally developed to solve the standard deconvolution problem, has undergone some minor modifications so that it can be used to assist the iterative phase retrieval algorithm in obtaining a solution to the phase retrieval problem. As with Parker and Bones (1992) point zero-matching algorithm, the phase retrieval problem must first be converted into a standard deconvolution problem by using the HIO algorithm to obtain a preliminary image estimate $\tilde{f}(x,y)$. The consistent algorithm can then use $\tilde{f}(x,y)$,

in the role of $\hat{h}(x,y)$, and the contaminated auto-correlation, in the role of $g(x,y)$, to obtain an improved solution to the phase retrieval problem. It should be noted that, as with Parker *et al*'s algorithm, $\tilde{f}(x,y)$ and $i(x,y)$ are truncated to the correct extent before applying the zero-matching algorithm. An overview of the modified consistent algorithm, which can be used to assist the iterative phase retrieval algorithms in obtaining a solution to the phase retrieval problem, is now given. The consistent algorithm has been extensively described in §5.1 and as a consequence the following overview is primarily limited to describing the configuration of the phase retrieval version of the consistent algorithm.

The phase retrieval version of the consistent algorithm uses the zero tracks in a $Z\{\tilde{\mathcal{F}}(\zeta,\gamma)\}$ slice to determine which zero tracks in the $Z\{\mathcal{I}(\zeta,\gamma)\}$ slice pertain to $Z\{\mathcal{F}(\zeta,\gamma)\}$. Using the same recognition and separation scheme as that described in §5.2, the algorithm then removes those zero tracks corresponding to $\mathcal{F}(\zeta,\gamma)$ from each $\mathcal{I}(\zeta,\gamma)$ slice. Unlike the consistent algorithm, the phase retrieval version does not attempt to reconstruct $\hat{f}(x,y)$ from the remaining zero tracks of $\mathcal{I}(\zeta,\gamma)$, but uses those zero tracks which it removes from $Z\{\mathcal{I}(\zeta,\gamma)\}$ to obtain $\hat{f}(x,y)$, i.e. whereas the consistent deconvolution algorithm removes the psf tracks, the phase retrieval version removes the tracks that correspond to $f(x,y)$. Henceforth the phase retrieval version of the consistent algorithm is known as zero track matching (ZTM) to distinguish it from both the algorithm of Parker *et al*, which matches point zeros, and the author's own consistent deconvolution algorithm. ZTM must be able to automatically improve $\tilde{f}(x,y)$ without outside intervention and so a particular configuration of the consistent algorithm, which was most likely to automatically provide improved image estimates, had to be selected. To this end, the author decided that ZTM would recognise and separate zero tracks of $\mathcal{F}(\zeta,\gamma)$ in five slices from $Z\{\mathcal{I}(\zeta,\gamma)\}$, i.e. a multiple slice version of the algorithm was chosen (see §5.3.2). The five zero sheet slices of $\tilde{\mathcal{F}}(\zeta,\gamma)$ and $\mathcal{I}(\zeta,\gamma)$ which were used by ZTM to obtain the zero tracks of $\hat{\mathcal{F}}(\zeta,\gamma)$ were those corresponding to ρ = 0.7, 0.8, 0.9, 1.0, and 1.1. This set of slices were chosen as results in §5.3.2 indicate that slices corresponding to $\rho \approx 1.0$ tend to provide the best results.

Once the zero tracks corresponding to $\hat{\mathcal{F}}(\zeta,\gamma)$ have been removed from the five $Z\{\mathcal{I}(\zeta,\gamma)\}$ slices, the zero tracks of $\hat{\mathcal{F}}(\zeta,\gamma)$ are sampled by ZTM so that $\hat{f}(x,y)$ can be obtained by using the algebraic reconstruction technique. The coarse r_γ sampling technique (see §5.3.1) is used to sample the $\hat{\mathcal{F}}(\zeta,\gamma)$ zero tracks as the fine r_γ technique is computationally too expensive. ZTM therefore sets r_γ to 1.0, 1.25, and 1.5, reconstructs an $\hat{f}(x,y)$ from each of the resulting sets of equations, and then selects the $\hat{f}(x,y)$ derived from the set of equations with the smallest κ as the best image estimate. The image $\hat{f}(x,y)$ differs from $f(x,y)$ by the arbitrary complex constant in addition to the inevitable contamination. The complex constant $C\exp^{i\psi}$ does not need to be completely resolved in the phase retrieval situation because $\exp^{i\psi} f(x,y)$ is an image form of $f(x,y)$. Thus ZTM uses $\|F(u,v)\|$, which by Parseval's theorem gives $\|f(x,y)\|$ (see Table 2.1), to determine the value of C and to scale $\hat{f}(x,y)$ so that it has the correct energy. As with the point zero matching algorithm, $\hat{f}(x,y)$ is then refined by performing further iterations. The author's phase retrieval algorithm that incorporates ZTM further refines $\hat{f}(x,y)$ by performing twenty HIO iterations and then fifty ER iterations. It has been found that ZTM can indeed assist the convergence of the iterative phase retrieval algorithm and results demonstrating the algorithms success are presented in §6.2.1

6.2 PERFORMANCE OF THE ZERO TRACK MATCHING ALGORITHM

6.2.1 Accuracy of Reconstructions

To evaluate how successfully ZTM assists HIO, an experimental study of the zero track matching algorithm's performance has been conducted. If ZTM was able to assist the HIO algorithm's convergence the results from the study should show a drop in the true error ϵ_T (see §3.5.2.1) at the iteration at which zero track matching was applied or shortly thereafter. The set of original images used in this study were four 8×8 and four 16 ×16 pixel complex images whose real and imaginary pixel values were generated in a random manner. The power spectra of these images were computed and random contamination was added to the spectra to give the resulting spectra either a 60 dB, 50 dB, 40 dB , or 30 dB SNR. The complete set of images was processed by HIO and when ϵ_T for the image estimate fell below a predetermined level, ZTM was applied to the current HIO image estimate to obtain a new image estimate. The predetermined levels for ϵ_T were chosen as -10 dB, -20 dB, -30 dB, -40 dB and -50 dB (or whichever of these were encountered). After obtaining the new image estimate, twenty further iterations of HIO, followed by 50 iterations of ER were performed as described in §6.1.2.

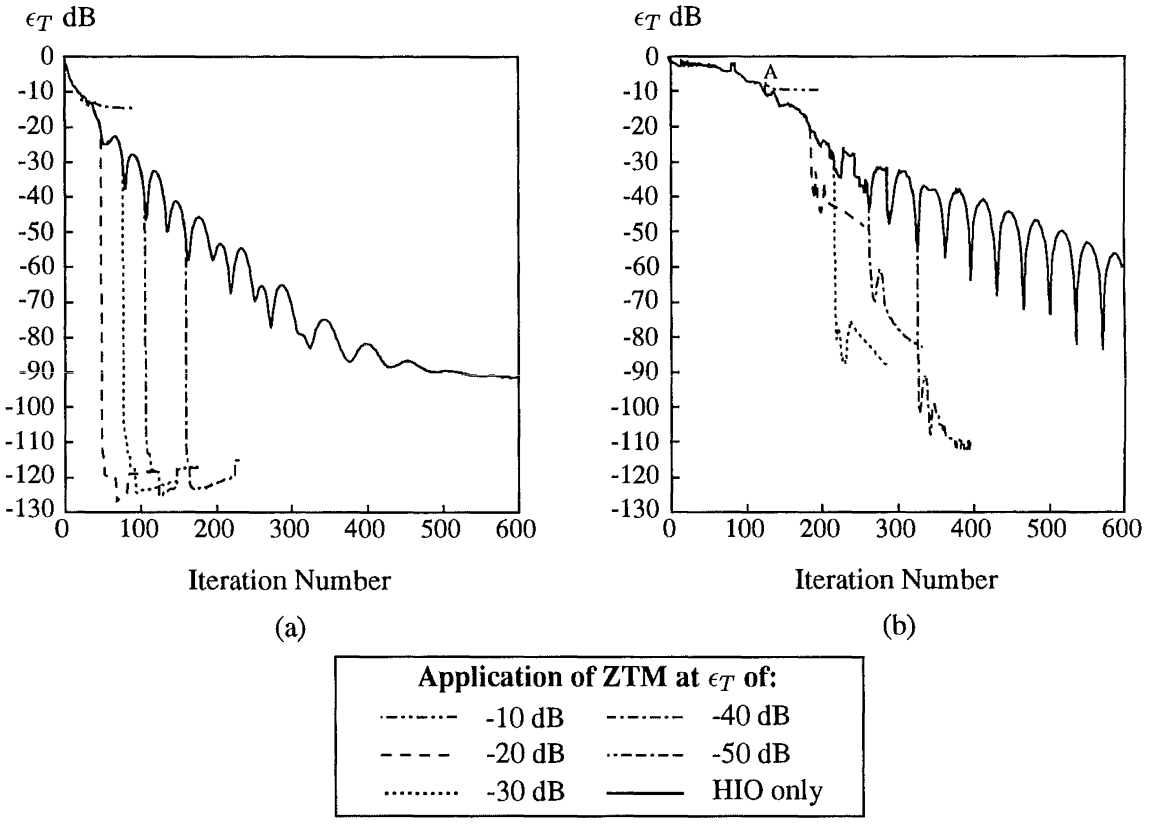

Figure 6.1 Plots of ϵ_T as a function of the number of HIO iterations with ZTM applied at various levels of ϵ_T for uncontaminated power spectra. (a) The results from an 8×8 $f(x, y)$, and (b) the results from a 16×16 $f(x, y)$.

In the situation where the power spectra were uncontaminated it was found that ZTM could make dramatic improvements to ϵ_T for the image estimate. This dramatic improvement is

shown in Figs. 6.1 (a) and (b), where the results obtained from using ZTM on a 15×15 pixel uncontaminated power spectrum and 31×31 pixel uncontaminated power spectrum are shown respectively. In these figures, ϵ_T is plotted as a function of the HIO iteration number and the results obtained on using ZTM are shown with broken lines, while those obtained from using HIO solely are shown as solid lines. For both these images the step decline in ϵ_T, at the iteration at which ZTM was applied, demonstrates that ZTM can dramatically assist the convergence of the HIO algorithm in the uncontaminated situation. The improvements in ϵ_T for the 16×16 original image are not as dramatic as those for the 8×8 image because, as explained in §5.4, the quality of the image estimates obtained from the zero-based algorithm tends to drop as the image size increases.

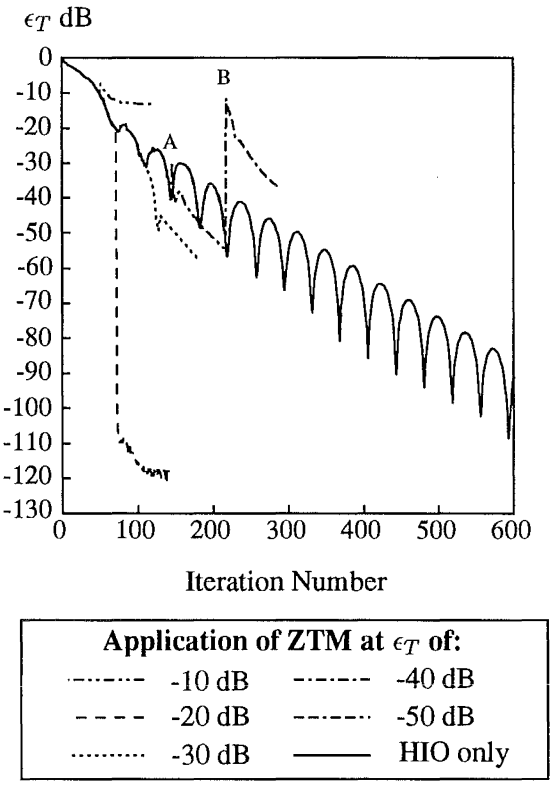

Figure 6.2 The ϵ_T *vs* HIO iteration number plot for a 16×16 $f(x,y)$ demonstrating that in some instances, with 16×16 $f(x,y)$, ZTM is unable to obtain improved image estimates.

While the results in Fig. 6.1 (b) are typical of those obtained in the uncontaminated case for a 16×16 pixel image, it was found on some occasions that ZTM would fail to obtain an improved image estimate. Results demonstrating this inability to obtain an improved image estimate are shown in Fig. 6.1 (b) at the point marked A and in Fig. 6.2 at the points marked A and B. The reason for ZTM's inability to obtain an improved image estimate in these situations is that the coarse r_γ sampling technique makes ZTM very susceptible to errors in the zero tracks of $\hat{\mathcal{F}}(\zeta,\gamma)$ arising because of the inaccuracy in $\tilde{f}(x,y)$. That is, the error in $\tilde{f}(x,y)$ means that ZTM is not always able to completely recognise all of the zero tracks of $\mathcal{F}(\zeta,\gamma)$ in the $Z\{\mathcal{I}(\zeta,\gamma)\}$ slice and, as the coarse r_γ technique only uses three different r_γ values, it is possible that an appreciable number of zero samples are always selected from the inaccurate regions of $Z\{\hat{\mathcal{F}}(\zeta,\gamma)\}$. In fact while investigating the consistency of the consistent algorithm (see §5.2) it was found that the Q *vs* r_γ plots would oscillate quite markedly for the larger images as the number of zero samples on these inaccurate zero track sections increased and

decreased with variation in r_γ. This effect is not observed for the smaller images (e.g. the 8×8 pixel image used to produce Fig. 6.1 (a)) because each $Z\{\mathcal{I}(\zeta,\gamma)\}$ slice for a smaller image comprises a smaller number of zero tracks and the separation algorithm therefore finds it easier to recognise those zero tracks of $\mathcal{F}(\zeta,\gamma)$ in the $Z\{\mathcal{I}(\zeta,\gamma)\}$ slice, i.e. the algorithm can tolerate more inaccuracy in the zero tracks of $\tilde{\mathcal{F}}(\zeta,\gamma)$ as there is a smaller number of zero tracks to choose from. It is anticipated that by either refining the coarse r_γ sampling technique, or by using a fine r_γ sampling technique, it should be possible to eliminate this problem with ZTM which arises for larger images. Furthermore, the results obtained from ZTM, which show a dramatic drop in ϵ_T, demonstrate that this algorithm has an enormous amount of potential for improving the convergence of the iterative phase retrieval algorithms.

In the situation where the power spectra were contaminated, it was found that ZTM could improve, quite significantly, the convergence of the HIO algorithm. Typical results obtained from one of the 15×15 pixel contaminated power spectra are shown in Figs. 6.3 (a), (b), (c), and (d) for SNRs of 60 dB, 50 dB, 40 dB, and 30 dB respectively. In Fig. 6.3 the quite noticable drops in the image estimates' ϵ_T, at the point where ZTM was applied, confirm that ZTM can indeed be used to improve the convergence of the HIO algorithm in the contaminated situation. This result is quite significant as it confirms the finding of Parker and Bones (1992) which shows that the set of power spectrum Z transform zeros can indeed be separated in the contaminated situation to give a set of zeros which provide a useful image estimate. The contamination on the power spectrum prevents the fall in ϵ_T being as dramatic as that noted in Fig. 6.1 for the uncontaminated case, nonetheless the decline in ϵ_T, with the application of ZTM is still appreciable. The results in Fig. 6.3 demonstrate that, for an 8×8 image, ZTM can have a beneficial effect on HIO convergence for power spectrum SNRs as low as 30 dB. In fact ZTM improved the convergence of HIO in three out of the four cases with 30 dB SNRs. The result for the 30 dB SNR power spectrum in Fig. 6.3 (d) is interesting in its own right as it shows that ZTM can in some instances provide image estimates of a quality which are likely to be superior to those obtained by HIO alone. The plot in Fig. 6.3 (a), obtained from the 60 dB SNR power spectrum, demonstrates that ZTM can also break the HIO algorithm out of stagnation, as ϵ_T remains reasonably stable for about 400 iterations when HIO is used alone but when ZTM is applied, at about 200 iterations, ϵ_T drops dramatically. This results confirms the finding of Parker and Bones (1992) that zero matching algorithms can break the HIO algorithm out of stagnation. Another interesting result that can be seen by inspecting the graphs in Fig. 6.3 is that ZTM can have a beneficial effect when the image estimate obtained from HIO has a SNR as low as 10 dB. As ϵ_V tends to track ϵ_T the above result suggests that in practice ϵ_V could be tracked and when it was noted that ϵ_V had fallen below a predetermined level, say -10 dB or lower, ZTM could be applied with a reasonable level of confidence.

An extensive set of results has only been presented for the 15×15 pixel contaminated power spectra as those obtained from the 31×31 pixel power spectra, in a large number of cases, suffered from the previously described image estimate sensitivity to the coarse r_γ sampling technique. Nonetheless, a graph of results obtained from applying the zero track matching algorithm to a 60 dB SNR 31×31 pixel contaminated power spectrum are shown on Fig. 6.4. The plot in Fig. 6.4 demonstrates that ZTM, in some situations, can successfully assist the convergence of the HIO algorithm when the original image is larger than 8×8 pixels and the power spectrum is contaminated. The fragile nature of ZTM with regards to the size of the original image made it difficult to determine how the algorithm could be expected to perform with larger images. It was noted, however, that there did appear to be a relationship between original image size and the size of the improvements that could be expected with a reliable ZTM. The size of the improvements obtained from the 31×31 pixel contaminated

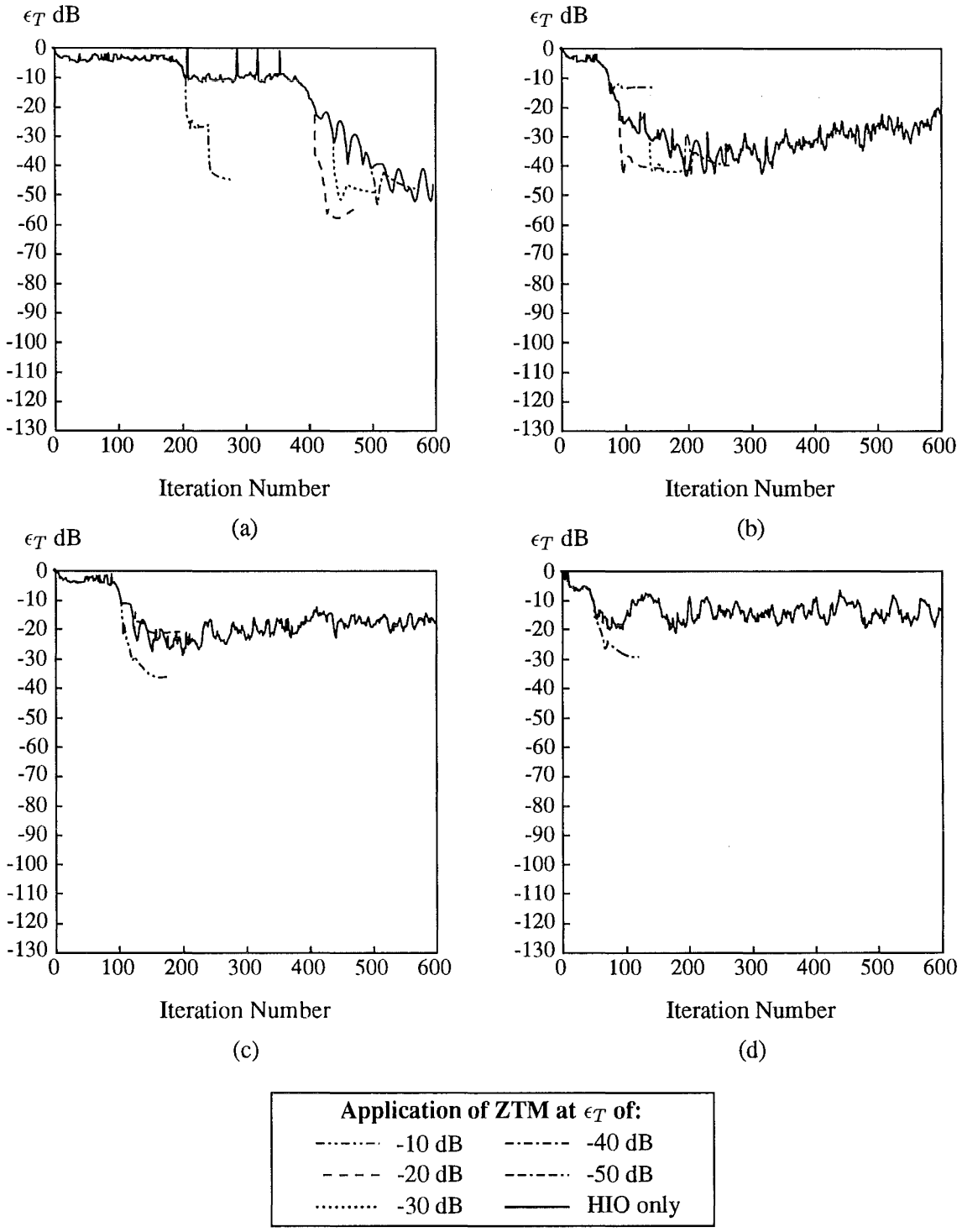

Figure 6.3 The ϵ_T *vs* HIO iteration number plots for contaminated 15×15 pixel power spectra with and without application of ZTM. The power spectra had SNRs of (a) 60 dB, (b) 50 dB, (c) 40 dB, and (d) 30 dB.

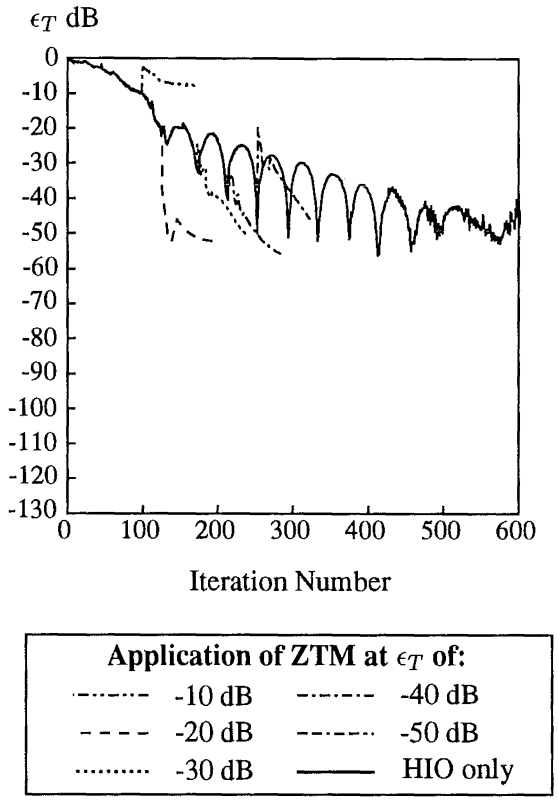

Figure 6.4 The ϵ_T *vs* HIO iteration number plot for a 60 dB SNR 31×31 pixel power spectrum with and without application of ZTM.

power spectra tended to be smaller than those obtained from the 15×15 pixel power spectrum. In some cases ZTM produced image estimates with ϵ_Ts greater than those of the image estimates originally provided by the HIO algorithm. The decline in the performance of ZTM as image size increases or as power spectrum (convolution) SNR decreases is to be expected due to the results reported in §5.4 for the deconvolution problem. In §5.4 the reasons for this decline in image estimate quality are explained. This decline in image estimate quality means that as the image size increases or the power spectrum SNR decreases ZTM finds it increasingly difficult to provide an image estimate with an ϵ_T that is lower than that originally provided by the HIO algorithm. It is interesting to note that ZTM was completely unable to assist the convergence of the HIO algorithm in recovering the phase of any of the 16×16 pixel images from the 30 dB SNR 31×31 pixel power spectra. This result should be compared to that reported above for the 8×8 images.

In summary, the results have shown that ZTM can assist the convergence of the HIO algorithm in obtaining a solution to the contaminated phase retrieval problem. ZTM can be successfully applied when the HIO image estimates have ϵ_Ts as low as -10 dB and when the power spectra have SNRs as low as 30 dB. As the size of the original image increases the performance of ZTM becomes increasingly fragile, although the results are still promising. The results also show that ZTM has the beneficial property that it can break the HIO algorithm out of stagnation. To determine whether ZTM can actually speed up the convergence of the HIO algorithm the computational expense of ZTM in comparison to HIO must be determined. A discussion on the computational expense of ZTM is given in §6.2.2. It is worth noting that the results reported here also demonstrate that the consistent algorithm could be used to solve the standard deconvolution problem when the component images are complex.

6.2.2 Computational Requirements

To determine whether ZTM could be used to speed up the convergence of the HIO algorithm, an experimental comparison of the computational requirements of both algorithms has been performed. To compare the computational expense of the algorithms, ten 4×4, ten 8×8, and ten 16×16 pixel complex random images had their phase retrieved from their power spectra by the HIO algorithm in conjunction with ZTM. The time required to complete the HIO iterations and the zero track matching process, for all the images, was then recorded to allow a comparison to be made. For the HIO algorithm the computation time was assessed for the situations where ϵ_V was and was not computed at each iteration, as ϵ_V may be used by the algorithm incorporating ZTM to indicate when ZTM should be applied (see §6.2.1). It was anticipated that ZTM would be computationally very expensive as for a $N \times N$ image the zeros of $5 \times 1024 = 5120$ $4N-4$ degree and 5120 $2N-2$ degree 1-D polynomials must be calculated to form the zero tracks comprising the five slices of the auto-correlation's and image estimates zero sheet respectively (see §5.1). This should be compared to the point zero matching algorithm (see §6.1.1) which requires the computation of the zeros of 2N+2 1-D polynomials of degree $\leq 4N-4$. In addition to the enormous computational cost of computing the zero tracks ZTM also uses the computationally expensive algebraic reconstruction technique to reconstruct $\hat{f}(x,y)$ from its zeros. It was shown in §5.6 that for images with an extent of about 16×16 pixels the algebraic reconstruction process becomes computationally more expensive than computing the zero tracks as the algebraic reconstruction technique's computational expense is $O(N^6)$. The process of matching the zero tracks is relatively inexpensive as was shown in §5.6.

The results from the comparison of the computational expense of HIO to that of ZTM are shown in Table 6.1. In Table 6.1 the number of HIO iterations that can be performed in the mean time required to perform one application of ZTM is given for each image size. The mean times referred to in this comparison are the means of the ten computation times obtained for the ten different images of the various sizes. These results confirm the anticipated result that ZTM is computationally very expensive in comparison to HIO. Table 6.1 shows that for a 16×16 image one application of zero track matching is equivalent to 44032 iterations of HIO in the situation where ϵ_V was not computed at each HIO iteration. Even for the smaller 4×4 image one application of ZTM was equivalent to 15673 HIO iterations. For any size of image the HIO algorithm in significantly less than 15673 iterations should be able to produce an image estimate with a quality that is superior, or at least equivalent to that provided by ZTM. The only situation where it is possible that ZTM could obtain a reconstruction superior to that provided in 15673 iterations by HIO alone, is that where the HIO algorithm had stagnated. The above belief is confirmed by the plots in Fig. 6.3 which show that, in the practical situation where the power spectra are contaminated, in less than 600 iterations the HIO algorithm is likely to provide an image estimate with a quality equivalent to that provided by ZTM.

COMPUTATION COST OF THE ZERO TRACK MATCHING ALGORITHM
IN TERMS OF HIO ITERATIONS

Image Size	ϵ_V computed	ϵ_V not computed
4×4	5361	15673
8×8	12370	28431
16×16	24530	44032

Table 6.1

The above situation is not going to improve for larger images as it can be seen in Table 6.1 that the computational expense of the zero track matching algorithm grows significantly

faster with increasing image size than that of the HIO algorithm. In fact for large images the ratio of the computational expense of ZTM to that of the HIO algorithm should be $O(N)^5$ as for large images the zero track matching algorithm's computational expense is $O(N)^6$ and the HIO algorithm's computational expense scales something like $O(N\log(N))$. The results in Table 6.1 demonstrate that the above findings are also true for the case where ϵ_V is not computed at each iteration. It is worth noting that for very small images (i.e. 4×4 images) the computation cost of ZTM could be reduced by reducing the number of zeros computed on each zero track. Care would have to be taken, however, to ensure that the point zeros could still be unambiguously linked to give the zero tracks. For larger images (i.e. 16×16) reducing the number of zeros would have a negligible effect on the computational cost as for images of this size the algebraic reconstruction technique is the major contributor to the overall computational cost. The author, at present, does not know of any method which could be used to reduce the computation cost of the algebraic reconstruction technique.

The above set of results have demonstrated that in most circumstances ZTM cannot speed up the convergence of the HIO algorithm. The dramatic improvement in the quality of the image estimates noted with ZTM comes at a great computational expense which is equivalent to many HIO iterations. It is therefore computationally more efficient to use HIO alone to obtain an image estimate of a required accuracy than to use HIO in conjunction with ZTM. The only situations where applying ZTM may be beneficial are those where the HIO algorithm has stagnated or where ZTM can provide an image estimate superior to that which can be obtained by HIO alone. It was shown in §6.2.1 for low SNR power spectra (i.e. 30 dB) and small images that ZTM can indeed provide image estimates of a quality superior to those provided by HIO alone (see Fig. 6.3 (d)).

6.2.3 Comparison to Point Zero Matching Algorithm

The performance of ZTM has been compared to that of the point zero matching algorithm (see §6.1.1) developed by Parker and Bones (1992). This comparison was conducted as it was believed that ZTM could provide image estimates of a superior quality to those provided by the point zero matching algorithm in the situation where the image estimate provided by the HIO algorithm is of poor quality. The reason why it was believed that ZTM could provide superior $\hat{f}(x,y)$ was that the algorithm attempts to recognise the complete zero tracks which allows the directional information contained within the tracks to be uses to assist in recognition, i.e. the fact that the zero tracks of the image estimate provided by HIO head in approximately the same direction as those in $Z\{\mathcal{I}(\zeta,\gamma)\}$ which belong to $Z\{\mathcal{F}(\zeta,\gamma)\}$ allows the tracks belonging to $Z\{\mathcal{F}(\zeta,\gamma)\}$ to be recognised. The point zero matching algorithm cannot use this directional information as it only matches point zeros. ZTM is computationally much more expensive than the point zero matching algorithm as the former algorithm needs the zero tracks of $\mathcal{I}(\zeta,\gamma)$ and $\tilde{\mathcal{F}}(\zeta,\gamma)$ to perform the zero matching process whereas the latter algorithm only computes one point zero from each track. ZTM also uses the algebraic reconstruction technique in contrast to the computationally inexpensive Fourier reconstruction technique. Unless ZTM is able to provide $\hat{f}(x,y)$ of a quality superior to those provided by the point zero matching algorithm, the enormous computational cost of ZTM would be unjustified.

To compare ZTM to the point zero matching algorithm the sets of 8×8 and 16×16 pixel complex random images described in §6.2.1 had their phase retrieved from their contaminated power spectra by the two different algorithms. The zero matching process, as described in §6.2.1, was implemented when the ϵ_T of the image estimate provided by HIO first fell below either -10 dB, -20 dB, -30 dB, -40 dB, or -50 dB. The image estimates were then compared after

twenty further iterations of HIO and fifty iterations of ER to see which algorithm provided the best image estimate improvement. Also compared in this evaluation were the smallest values of ϵ_T at which zero matching provided an improved image estimate, the computational cost of the two algorithms, and the effects of increasing image size on the image estimate quality provided by the two techniques.

COMPARISON OF THE ZTM AND POINT ZERO MATCHING ALGORITHM FOR VARIOUS POWER SPECTRUM SNRs

	8×8 $f(x,y)$			16×16 $f(x,y)$		
$I(u,v)$ SNR	ZTM Q dB	PZM Q dB	Δ Q dB	ZTM Q dB	PZM Q dB	Δ Q dB
∞	94.34	80.63	12.71	50.43	68.97	-18.54
60	48.80	47.60	1.2	39.45	44.34	-4.89
50	39.19	40.29	-1.1	26.50	34.21	-7.71
40	32.31	30.46	1.85	17.40	27.76	-10.36
30	23.57	20.83	2.74	8.43	15.33	-6.90

Table 6.2 ZTM = Zero track matching, PZM = Point zero matching, and ΔQ = ZTM Q dB - PZM Q dB.

In Table 6.2 the results obtained from the two algorithms are tabulated with respect to the various power spectrum SNRs. The Q values stated are the mean of all the Q values obtained when zero matching was applied at the various values of ϵ_T for the four different images. The results from the 8×8 image show that ZTM on average provides an improved image estimate of a quality superior to that provided by the point zero matching algorithm, although the differences between the quality of the reconstructions, for the contaminated power spectra, are quite small. For the 16×16 pixel image the point zero matching algorithm always provides improved image estimates of a quality superior to those provided by ZTM. In Table 6.3 the mean reconstruction qualities obtained from the four different 60 dB SNR power spectra of each size are tabulated against the various values of ϵ_T at which zero matching was applied. The sets of results for the different size images both demonstrate that ZTM can provide improved image estimates of a superior quality when the image estimate provided by HIO is poor, i.e. when ZTM is applied at ϵ_T = -10 dB the improved image estimate obtained is superior to that obtained from point zero matching. The major advantage that the point zero matching algorithm has over ZTM is its relatively small computational cost. For the 8×8 image one application of point zero matching was equivalent to 53 HIO iterations, whereas one application of ZTM is equivalent to 28431 HIO iterations. For the 16×16 image the computational cost of the point zero matching algorithm and ZTM are equivalent to 65 and 44032 iterations of HIO respectively.

The above results demonstrate that the belief that ZTM should provide superior improved image estimates, in the situation where the image estimate provided by HIO is of a poor quality, was correct. It appears that ZTM is able to use to advantage the directional information contained within the zero tracks of the image estimate to obtain a more accurate estimate of the zero tracks of the image and an improved image estimate. This superior performance, however, comes at an enormous computational cost and it is questionable whether the small superiority of ZTM, as compared to the point zero matching algorithm, could be used to justify the cost of ZTM. When the image size becomes large the difficulty in obtaining an accurate well-conditioned set of equations, for the reconstruction algorithm, again appears to limit the performance of ZTM. For the 16×16 image it was noted that the point zero matching algorithm never degraded the quality of the image estimate provide by HIO, which is in contrast to the

COMPARISON OF THE ZTM AND POINT ZERO MATCHING ALGORITHM AT DIFFERENT VALUES OF ϵ_T

	8×8 $f(x,y)$			16×16 $f(x,y)$		
ϵ_T dB	ZTM Q dB	PZM Q dB	Δ Q dB	ZTM Q dB	PZM Q dB	Δ Q dB
-10	40.44	32.62	7.82	15.40	14.61	0.79
-20	52.47	52.29	0.18	48.87	47.81	1.06
-30	50.94	52.78	-1.84	45.90	52.95	-7.05
-40	51.30	52.70	-1.40	40.63	53.18	-12.55

Table 6.3 ZTM = Zero track matching, PZM = Point zero matching, and ΔQ = ZTM Q dB - PZM Q dB. The $I(u,v)$ SNR was 60 dB

results reported in §6.2.1 for ZTM. The above results should also be compared to those reported in §5.8 where a comparison was made between the zero track matching and point zero matching algorithms developed to solve the deconvolution problem. The results in §5.8 indicated that point zero matching was superior to zero track matching, however the results here suggest that the relatively high SNRs of the psfs used in §5.8 could be responsible for this result, i.e. the psfs were not sufficiently inaccurate for the superiority of zero track matching to be evident. The results presented here also confirm the author's belief that it is the reconstruction section of the zero track matching algorithms which is limiting the algorithms' performance and not the zero track recognition section of the algorithms.

CHAPTER 7

CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH - PART I

In Part One of this thesis a new algorithm which can successfully solve the contaminated 2-D standard deconvolution problem is described. The algorithm deconvolves the blurred image by separating the two component zero sheets, which together form the zero sheet of the blurred image. This separation process is performed by using the zero sheet of the psf estimate to recognise and remove the convolution zero sheet that belongs to the psf actually responsible for the blurring.

In Chapter 2, the mathematical models that describe the image formation and image degradation processes are presented. These models allow an expression to be derived which describes blurred and contaminated images. The existing techniques that attempt to invert this expression to obtain an estimate of the actual image, when the psf is known *a priori*, are described in Chapter 3. It is shown that the accuracy of these techniques is likely to be dependent on the accuracy of the psf estimate and it is also shown that the psf estimate is likely to be inaccurate. By using the zero sheet to deconvolve an image it is possible to reduce this dependence on psf accuracy, and so the concept of the zero sheet and some of its useful properties are introduced in Chapter 4. The new algorithm that solves the deconvolution problem by separating the convolution component zero sheets is described in Chapter 5 and results demonstrating the strengths and weaknesses of the algorithm are presented. A similar problem to deconvolution is phase retrieval and so the new algorithm has also been adapted to solve this problem. The phase retrieval version of the new algorithm is described in Chapter 6 and results demonstrating the success of the algorithm in this area of application are presented. The original work in Part One is predominantly contained within Chapters 5 and 6, although the research related to some of the properties of the zero sheet described in Chapter 4, particularly those related to the various effects of contamination, is also original.

The purpose of this Chapter is to present the conclusions that have been drawn from the results presented in Chapters 4,5, and 6 and to put forward the possible avenues of further research suggested by these conclusions. In §7.1 these conclusions are reported and the most important aspects of the work presented in Part One are summarised. The discussions in Chapters 4,5, and 6 are reasonably extensive and so the scope of the summaries is limited to the results that have a direct bearing on the conclusions made. In §7.2 the suggestions for further work are presented.

7.1 CONCLUSIONS

The results presented in §5.2 and §5.3 demonstrate that the new zero-based deconvolution algorithm, described in §5.1, can successfully solve the contaminated 2-D standard deconvo-

lution problem. The algorithm uses the consistency condition of Bates *et al.* (1976) to separate the component zero sheets and, as the results in §5.2 show, the algorithm is able to obtain consistent solutions to the deconvolution problem in the absence of contamination, unlike traditional deconvolution techniques. In the situation where the convolution is contaminated, the algorithm is also able to successfully use the consistency condition to separate the component zero sheets and obtain useful image estimates (see §5.2). Unfortunately, the algorithm in its present form can only obtain useful image estimates from small, high SNR convolutions (i.e. original image sizes $\leq 16\times16$ pixels and SNRs ≥ 40 dB, see §5.4). Nonetheless, this algorithm is the first algorithm to successfully separate the zero sheets of a contaminated convolution, in a completely automatic manner, to obtain a solution to the contaminated 2-D deconvolution problem.

On evaluating the sensitivity of the image estimate quality to the image size and convolution SNR, it was found that quality decreased whenever there was a corresponding increase in image size or decrease in SNR. This decrease in image estimate quality occurs because the likelihood of obtaining an accurate well-conditioned set of equations, for the algebraic reconstruction technique, is directly dependent on the above factors. When the SNR decreases or the original image size increases, it becomes more difficult to obtain a sufficient number of linearly independent equations to allow a useful solution to be obtained. Thus, the reconstruction technique appears to be the section of the algorithm that is currently limiting the useful application of the algorithm to practical problems. From another viewpoint, the enormous computational expense of the current reconstruction technique is also limiting the practical application of the algorithm. The evaluation of the computational expense of the algorithm, in §5.6, demonstrated that the algebraic reconstruction technique is computationally very expensive and, if original images larger than 16×16 pixels are to be reconstructed, may be prohibitive.

One of the main factors which motivated the development of this algorithm was the expectation that the accuracy of the image estimates provided would be less dependent on the psf accuracy than those provided by existing techniques. The results presented in §5.7 demonstrate that the image estimates provided by the consistent algorithm have an accuracy which is reasonably independent of the psf accuracy. The accuracy of the solution obtained only significantly declines when the psf becomes so inaccurate that its zero sheet can no longer be used to recognise the zero sheet of the actual psf. This is in contrast to the Wiener filter, which produces image estimates of a quality that steadily decline with the accuracy of the psf. The robustness of the consistent algorithm, with respect to psf inaccuracy, occurs because the algorithm only uses the psf estimate to recognise the actual psf and, unlike the Wiener filter, does not explicitly use the psf estimate to calculate the image estimate. The results also show, however, that as the original image size becomes large ($> 16\times16$ pixels) or the SNR of the convolution becomes small (≤ 40 dB), the quality of the image estimates provided by the Wiener filter are so much better than those provided by the consistent algorithm, that the sensitivity with respect to psf accuracy is of little consequence. These results demonstrate that the most useful application for the consistent algorithm, at present, is deblurring small, high SNR blurred images when the psf is inaccurate, where it performs much better than the Wiener filter. This situation is therefore the only situation where the enormous computational cost of the consistent algorithm can be justified in terms of a beneficial outcome.

It is shown in §4.4 and §5.3.3 that the bandwidth of the contamination on an image can have a marked effect on the manner in which the zero sheet of the image is distorted by the contamination. As explained in §4.4.1, when the contamination is wide-band the entire zero sheet is distorted in a manner that makes it impossible to predict whether different sections of the zero sheet have been distorted more, in relative terms, than other sections. In contrast,

when the contamination is bandlimited, specific regions of the zero sheet are distorted more than others and, if the contamination bandwidth is known *a priori*, it should be possible to predict which regions of the zero sheet exhibit less distortion. The results presented in §5.3.3 confirm that in the bandlimited case it is possible to improve the quality of the image estimate, by weighting the equations produced from the zero samples taken from the uncontaminated regions of the zero sheet heavier than those from the remaining regions. The results suggest that this prior knowledge cannot be used to dramatically improve the quality of the reconstructions, but it can refine a reconstruction that is already of a reasonable quality.

Adapting the consistent algorithm so that it could assist an iterative phase retrieval algorithm, it was found that one application of the zero-based algorithm could dramatically decrease the error on the image estimate (see §6.2.1). As shown in §6.2.2, this dramatic drop in image estimate error comes at an enormous computational cost (i.e. a single application of the zero based algorithm is equivalent to 44032 iterations of hybrid input output) and so the the zero-based algorithm was unable to successfully speed up the convergence of the iterative algorithm. It was also found with the larger 16×16 image that the zero-based algorithm was not very robust and in some cases increased the error on the image estimate. As explained in §6.2.1, the explanation is that as the image size becomes large, the zero-based algorithm finds it more difficult to provide an accurate well-conditioned set of equations that characterise the image. Nonetheless, for the smaller images (8×8 pixels) it was found that the zero-based algorithm could in some instances provide image estimates of a quality superior to those obtained by iteration alone and could break the iterative algorithm out of stagnation. Thus in at least these two situations the application of the zero-based algorithm could have a beneficial outcome.

7.2 SUGGESTIONS FOR FURTHER WORK

It is stated above that the algebraic reconstruction technique, with its ill-conditioned equations and large computational cost, is the section of the algorithm which is currently believed to be limiting the consistent algorithm's performance. To improve the quality of the image estimates obtained from the algorithm, the author believes that the following aspects of the reconstruction process would be worthy of further work. The first area is the implementation of a modified version of the coarse r_γ sampling techniques described in §5.3. In this modified version, r_γ would be set to a preset number of values between 1.0 and 1.5 (probably more than the three values used with the existing coarse r_γ technique) and the image estimates obtained from each value of r_γ would be averaged. As explained in §5.3, a least squares technique must be devised to equate all the complex constants that scale each of the image estimates. It would then be necessary to investigate whether there exists an optimum value for the size of the increment in r_γ, which provides average image estimates of a quality similar to those obtained from the fine r_γ technique also described in §5.3. While this technique may be able to reduce the amount of computation required to obtain a useful solution, the technique is still going to be computationally expensive, in comparison to the Fourier reconstruction technique, and is unlikely to improve the quality of the image estimates as either the image size increases or the convolution SNR decreases. It therefore seems likely that it may be necessary to investigate other techniques that allow an image to be reconstructed from its zeros.

Another reconstruction technique, which may improve the quality of the image estimates, is to use singular value decomposition (SVD), instead of QR decomposition, in conjunction with the algebraic reconstruction technique. When an overdetermined system of equations must be solved in a least squares sense and the system of equations is known to be ill-conditioned, a

solution superior to that which can be obtained by other linear equation solution techniques can often be obtained using SVD. In fact Press *et al.* (1986) note that SVD is the method of choice for solving the linear least-squares problem. Thus, it is possible that SVD may improve the quality of the image estimates obtained by algebraic reconstruction, when the conditioning of the system of equations decreases because of an increase in image size or decrease in convolution SNR. With SVD the $M \times N$ matrix $\mathbf{A}$ representing the coefficients of a system of equations, $\mathbf{Ax} = \mathbf{b}$, is decomposed into the form

$$\mathbf{A} = \mathbf{UWV}^T,$$

where $\mathbf{U}$ is a $M \times N$ column-orthogonal matrix, $\mathbf{W}$ is an $N \times N$ diagonal matrix with positive or zero (singular) values and $\mathbf{V}$ is a $N \times N$ orthogonal matrix (Dongarra *et al.*, 1979; Press *et al.*, 1986). This decomposition, as with QR decomposition, is carried out with Householder transformations (see §4.5.2.2). As $\mathbf{A}^{-1} = \mathbf{V}[\text{diag } 1/\mathbf{W}_{nn}]\mathbf{U}^T$, where $[\text{diag } 1/\mathbf{W}_{nn}]$ is the diagonal $N \times N$ matrix whose elements are the reciprocals of the $\mathbf{W}_{nn}$ comprising $\mathbf{W}$, the solution

$$\mathbf{x} = \mathbf{V}[\text{diag } 1/\mathbf{W}_{nn}]\mathbf{U}^T\mathbf{b}, \tag{7.1}$$

can be obtained by a series of simple multiplications. If the set of equations are ill-conditioned, some of the $\mathbf{W}_{nn} \approx 0$. When this occurs, the corresponding $1/\mathbf{W}_{nn}$ in Eqn.(7.1) is set to zero to remove the effect of these ill-conditioned equations from the least squares solution (Press *et al.*, 1986). It has been found that performing the above operation can often reduce the error on the solution quite substantially (Press *et al.*, 1986). In addition SVD, usually provides some indication as to why the traditional linear equation solution techniques are unable to provide a satisfactory solution. While it is possible that SVD may be able to improve the quality of the image estimates obtained from the consistent algorithm, this improvement comes at an enormous computational expense. As the image size becomes large, SVD is approximately three times more computationally expensive than QR decomposition and, as for QR decomposition, the computational cost is $O(N^6)$ for an $N \times N$ image. It would therefore by interesting to determine whether SVD could improve the quality of the image estimates, but its computational cost is probably going to prevent its practical application for large images.

The final area of further work in the reconstruction area, possibly leading to an improvement in image estimate quality, is to use the Fourier reconstruction technique in conjunction with an improved recognition and separation technique. The results in §5.5 demonstrated that the current recognition and separation technique cannot provide an image estimate zero sheet of an accuracy sufficient to allow Fourier reconstruction to be used. The improved technique may use some of the suggestions outlined in the following paragraphs to obtain an improved estimate of the image's zero sheet and an interpolation scheme could possibly be used to fill in any missing sections of the zero tracks. If it is possible to obtain an improved set of zero samples for the Fourier reconstruction technique, this improvement should translate into an improvement in the image estimate quality. This avenue of further research is probably the most promising area and the low computational cost of Fourier reconstruction, in comparison to the algebraic techniques, indicates that it is the most practical solution.

In the area of improving the estimate of the image's zero sheet, there are two possible areas of further research which could result in improvements in this area. The first is to use the properties of symmetry, outlined in §4.3.3, which the zero sheets of certain images possess, to improve the quality of the image estimate's zero sheet. For example, when an image is blurred by out-of-focus blur, it is known that the zero tracks of the psf at ρ and $1/\rho$ should be the same and that the zero tracks should also have symmetry through the unit circle. This information could therefore be used to assist the recognition and separation algorithms in obtaining a more

accurate estimate of the image's zero sheet. The second technique is to use the Wiener filter to obtain an estimate of the image and then use the zero sheet belonging to this estimate, in conjunction with the zero sheet of the psf estimate, to obtain from the convolution zero sheet a more accurate estimate of the zero sheet belonging to the image.

It is also possible that the consistent algorithm may be able to assist the Wiener filter in obtaining a good quality image estimate when the psf estimate is inaccurate. If the convolution has a high SNR, the image is small, and the psf estimate is inaccurate, it is very likely that the consistent algorithm could be used at first to obtain an improved psf estimate. For example, the results in Table 5.10 demonstrate that for a 60 dB SNR 9×9 convolution the consistent algorithm can obtain an image estimate with a Q of 31.5 dB when the psf has a SNR of 20 dB. It is anticipated that reconstructions of the same quality would be obtained if the psf was reconstructed from its zero sheet using the consistent algorithm. This improved psf estimate could then be used by the Wiener filter to obtain an improved image estimate. It may also be possible to combine the Wiener filter and consistent algorithm in an iterative loop to obtain an image estimate of optimal accuracy. If this was done, it would probably be most beneficial to use the point zero matching algorithm of Parker (1992), as the computation expense of the consistent algorithm is likely to be prohibitive.

The final suggestion for further work is to investigate the effects of convolution truncation on the quality of the image estimates provided by the zero-based techniques. It is anticipated by the author that, if the truncation is not too severe, the edge extension method of McDonnell and Bates (1975) may allow the consistent algorithm to reconstruct an untruncated version of the image estimate. This is because if the truncation is not severe, edge extending the convolution is going to give the convolution the correct extent without adding a significant amount of energy to its spectrum. A reasonable estimate of the zero sheet of the image may then be able to be recognised within the zero sheet of the extended convolution. It would also be interesting to compare the quality of the image estimates obtained from the consistent algorithm to those obtained from the Wiener filter when the convolution is truncated. It is possible that the consistent algorithm may simply provide a reasonably good estimate of a truncated version of the image.

Part II

Evaluation of Human Sensory-motor Function

CHAPTER 8

THE HUMAN SENSORY-MOTOR SYSTEM AND ITS EVALUATION

For many centuries human beings have been interested in gaining an understanding of the function and dysfunction of the various systems which comprise the human body. This interest is most likely motivated by the self preservation instinct, a desire to improve the quality of life of those fellow human beings whose body has begun to dysfunction for any number of reasons, and the intrinsic quest of human beings for knowledge on themselves and the world around them. One of the systems within the human body, which has received considerable attention over the last century, is the sensory-motor (S-M) system. The S-M system is the system responsible for controlling the movement of the human body's extremities in a coordinated and voluntary manner, allowing some predefined goal to be achieved. The S-M system comprises the brain, the sensory organs, the limbs, and those nerves and components of the musculoskeletal system which are associated with controlling the movement of the limbs.

Much of the interest in the function of the S-M system stems from the fact that many neurological disorders disrupt the integrity of the S-M system. These disruptions often prevent or limit the ability of a person to perform activities of daily living. In addition, to enable people to achieve previously unobtainable goals and to improve the quality of life, people design many machines and apparatus. These machines are generally used in conjunction with a human operator and an understanding of the S-M system is therefore required if the human operator and machine are to be successfully integrated into one system. Until recent times, the performance of the S-M system in both its normal and impaired states has primarily been assessed by getting subjects to perform simple manual tasks. These tasks are designed so that the information required on the function of the S-M system can be inferred from the ease with which the tasks were completed or the strategy used by a subject to complete the task. The performance measures used with these tasks are generally subjective in nature and when quantitative measures are used it is generally not possible to gain quantitative information on the strategy used to complete the task. There was therefore a need to develop tests able to assess more complicated S-M function which could provide quantitative information on S-M performance and strategies. To satisfy this need researchers developed a series of tasks known as tracking tasks. Tracking tasks provide quantitative measures of S-M function by quantitatively assessing the ability of a subject to follow a moving visual target with a pointer controlled by the coordinated movement of a limb. Mechanical and electronic devices are used to provide quantitative measures of the subject's ability to accurately follow the target movement. Advances in computer technology over recent years have made it significantly easier to implement versatile tracking tasks and to quantitatively analyse the performance of the subject completing the task.

In Part Two of this thesis a set of 2-D tracking tasks (i.e. the target and subject's response

have two spatial degrees of freedom) developed by the author are described along with the original research conducted to evaluate the usefulness of these tasks. The set of 2-D tracking tasks have been developed to fulfil a need within the recently developed computerised S-M test battery at Christchurch Hospital (see Jones *et al.* (1993) and §8.4) for a series of tracking tasks that could evaluate integrated upper-limb sensory motor function in a multidimensional environment. The 2-D tracking tasks are more closely allied to the multidimensional environment in which the S-M system must function than the existing 1-D tracking tasks and, for this reason, it was considered that the 2-D tasks might be able to provide more valid information on S-M mechanisms than their 1-D counterparts. It was also considered that the 2-D tracking tasks might provide new or more sensitive measures of S-M function than the existing 1-D tracking tasks.

The set of 2-D tracking tasks comprise a random pursuit tracking task and a step pursuit tracking task. The 2-D random pursuit tracking task is not particularly novel as several other researchers have developed similar tasks (see §8.3.4). The 2-D step tracking task is however, at least to the author's knowledge, novel. The main thrust of the author's original research in this area was to determine whether these 2-D tracking tasks could indeed provide either new or additional information on S-M function which cannot be provided by 1-D tracking tasks. There was a need to determine whether there was any benefit to be gained from using 2-D tasks in preference to 1-D tasks, since the performance analysis required with 2-D tracking tasks is more complicated than that for 1-D tasks and 1-D tasks are generally easier to implement. The author also investigated whether the 2-D tasks could demonstrate the motor deficits known to occur in subjects with a common neurological disorder and which are clearly evident in their 1-D tracking performance. The final area of research was to establish why the adapted cross-correlation technique, within Christchurch Hospital's original S-M test battery, used to determine the time delay between the target signal and the subject's response signal, in some instances, produced quite inaccurate estimates of the time delay. To evaluate why the adapted cross-correlation technique provided inaccurate time delay estimates the author performed a mathematical analysis of the technique and compared the performance of this technique to other time delay estimation techniques in an experimental study using simulated response signals.

The reader is referred to §2.1 for the notation used in Part Two of this thesis. Also Part One, §2.4 and §2.5 provide the necessary mathematical preliminaries for signal processing procedures such as correlation and the Fourier transform.

In this chapter a brief overview of the S-M system is given and the techniques which can be used to evaluate the system's performance are described. In §8.1 the function of the various subsystems which comprise the human S-M system are described. In this section a simple model which describes the relationship between, and the function of, these subsystems is introduced. Subjects with Parkinson's disease were used to evaluate whether the 2-D tasks could be used to determine the S-M deficits known to arise with this disorder. For this reason the pathophysiology of Parkinson's diseases and the usual symptoms associated with this disease are described in §8.2. The techniques that can be used to evaluate the performance of the S-M system and investigate the mechanisms responsible for various neurological disorders are described in §8.3. In this section the traditional (largely subjective) neurological tests, quantitative psychomotor tests, tracking tasks in general, and 2-D tracking tasks are described. Finally, in §8.4 the computerised S-M test battery at Christchurch hospital is briefly described. This discussion includes the design philosophy, a description of the hardware, and an overview of the software. This computerised S-M test battery is the first comprehensive S-M test battery of upper-limb function to be implemented on a versatile computer platform such as a personal computer.

8.1 THE HUMAN SENSORY-MOTOR SYSTEM

One of the factors responsible for the survival of humanity is the ability to move in a purposeful manner. This ability to move allows one to feed, reproduce and, in life threatening situations, to either escape or fight. Most of the movements associated with self preservation are reflex or genetically programmed actions that do not have to be learnt, such as blinking an eye when an object unexpectedly approaches the eye. Humankind also has the ability to perform skilled or learned movements which allow the environment to be manipulated in a predetermined and efficient manner to achieve some beneficial outcome. The physiological system responsible for both these forms of movement, and which is therefore responsible for survival and evolution, is the S-M system. The S-M system uses the sense organs to gain information on the current state of the environment and then determines, usually in association with some predefined goals, an appropriate physical response to the environment's current state. Once the nature of the physical response has been determined, the S-M system executes the response by coordinating the movements of the musculoskeletal system. In this section the function of the major subsystems that comprise the S-M system are described. Also described is the manner in which these subsystems interact to allow both skilled and unskilled purposeful movements to be executed in response to environmental stimuli.

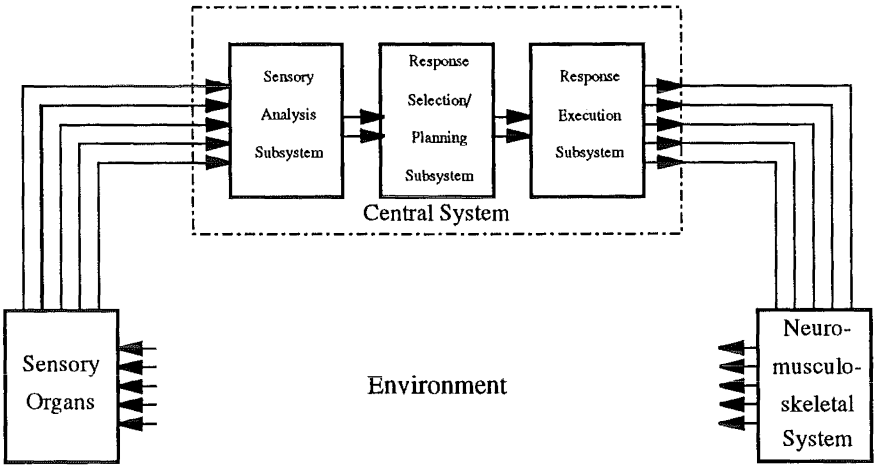

Figure 8.1 A model of the S-M system showing the physical components of the system (i.e. the sense organs, central nervous system, and limbs) and the information processing subsystems within the central nervous system which include the sensory analysis, response selection/planning, and response execution subsystems.

To allow the function of the S-M system to be described in a straightforward manner, it is useful to combine knowledge of the physical structure of the S-M system with the concept that the S-M system is an information processing system with a finite bandwidth, to produce a model of the system. It has been noted by Whiting (1975) and Schmidt (1988) that the concept of the S-M system being a bandlimited information processing system has proven to be very popular and useful in terms of gaining an understanding of the system. The hierarchical three-stage model commonly used to describe the S-M system is shown in Fig. 8.1 (Welford, 1968; Whiting, 1975; Schmidt, 1988). More complicated mathematical models of the S-M system, such as those of Neilson *et al.* (1992), follow the overall form of this model. At the top level of the model, which is the physical level, the model comprises the sense organs (i.e. the eyes, ears, nose, taste receptors, and tactile receptors), central system (i.e. the central nervous system (CNS) comprising the brain and spinal cord), and the motor components of the neuro-

musculoskeletal system (i.e. the peripheral nervous system, and muscles and bones of the limbs). The sense organs are the input to the S-M system and they provide information, such as visual, audio, and tactile information, on the current state of the environment. The sense organs in providing this information convey (or feedback) to the S-M system the effect that past movements had on the state of the environment. The information provided by the sense organs is passed to the central system which is the information processing system within the S-M system. According to the model, the central system consists of subsystems to carry out sensory analysis, response selection/planning, and response execution (Welford, 1968; Whiting, 1975; Schmidt, 1988). The function of each of these subsystems and the way in which they interact is described in detail later in this section. The function of the central system as a whole is to determine which sensory information is relevant, to decide on an appropriate response to the sensory information, to plan the movement, and to execute the movement. The neuro-musculoskeletal system is the output of the S-M system and consists of those components of the overall musculoskeletal system involved in performing coordinated movement. The appropriate muscle groups within the musculoskeletal system produce an output, in response to the stimuli initially provided by the sense organs, on being activated by signals travelling along the peripheral nervous system from the central system.

The three subsystems which form the central system are all multichannel and operate continuously and in parallel to carry out the information processing required by the S-M system. The belief that the systems operate in parallel is consistent with the fact that the three subsystems are physically located in different regions of the brain (Neilson *et al.*, 1992). The first subsystem to be described is the sensory analysis subsystem which accepts many channels of information on the current state of the environment from the sense organs. Much of the information provided by the sense organs has a low information content as it is of either no relevance to the task being performed or is redundant. Redundancy in the sensory information arises because many of the senses can provide many channels of the same information. For example, if a hand is rested on a desk, the hand's location may be conveyed to the brain via tactile sensory information channels and visually. Neilson *et al.* (1992) notes that this redundancy greatly improves the reliability of the transmission of information between the sensory receptors and the brain but would greatly increase the workload of the CNS if all the information was to be analysed. As alluded to before, the central system has a finite bandwidth and it is therefore unable to process all the information that it is continuously being provided by the sense organs. The sensory analysis subsystem therefore acts like an adaptive filter and only passes on, for further processing, the information which is unique and relevant to the task being performed.

If either the task currently being performed has been performed before or the stimuli being received have been received before, the sensory analysis subsystem recognises this and uses information contained within memory to determine what sensory information is relevant. Thus, with repeated exposure to a task, the S-M system learns what sensory information associated with a task is relevant and in future is able to focus solely on those sensory inputs which provide the most information, i.e. the concept of adaptive filtering or selective attention (Whiting, 1975; Neilson *et al.*, 1992). The need to process more information when a task is novel is one of the reasons why it is more difficult to perform novel skilled tasks than familiar skilled tasks. While filtering out the redundant information the sensory analysis subsystem is also processing the relevant sensory information and using memory to identify those features contained within the environment which it has had previous exposure to and has a high level description of. For example, the sensory analysis subsystem is able to identify a white, smooth, oval object as an egg (Neilson *et al.*, 1992). After filtering out the redundant sensory information and

processing the relevant information, to allow features in the environment to be identified, the sensory analysis subsystem provides a high level description of the environment to the response selection/planning subsystem.

After receiving information from the sensory analysis subsystem on the current state of the environment, the response selection/planning subsystem must plan an appropriate response. The first task of this subsystem is therefore to select the motor response which satisfies some predetermined goal and which is consistent with the current state of the environment. To select the appropriate response from the large set of possible responses the system uses memory of the effect of similar past responses to determine which response is most likely to have the desired outcome. Schmidt (1988) notes that some researchers postulate that the sensory analysis subsystem may also have a part to play in response selection as some sensory stimuli are believed to activate the memory of past responses associated with that stimuli. The use of memory indicates that the response selection/planning subsystem is able to learn which response is most appropriate and each time a particular response is performed it is refined and stored in memory for future reference. The ability of the S-M system to adapt its responses in this manner is one of the reasons why it is possible to become very efficient (skilled) at performing a particular task. Once the appropriate response has been chosen, Neilson *et al.* (1992) propose that the response selection/planning subsystem then divides the response up into a hierarchical structure of subresponses or subgoals. Each of these subresponses leads to the next subresponse and by concatenating the subresponses the overall response is obtained. For example, if the appropriate response in a particular situation is to catch a ball, at one level in the hierarchy the subresponses may be: run towards the ball, outstretch arms, close hands round ball, and move arms back towards the body. Once this hierarchy of subresponses has been determined, the response selection/planning subsystem produces a movement plan that describes how the motor system is to perform the next lowest level response (Neilson *et al.*, 1992). This subresponse is stored in memory for the response execution subsystem to use in producing the actual response. It is necessary to note that response selection/planning subsystem is continually evaluating the current state of the environment and if at any stage it is found that the overall response is no longer valid, implementation of the hierarchical structure of subresponses is terminated at the completion of the current lowest level subresponse. For example, in the ball example if the ball was to hit the ground before it could be caught, the overall response plan could be terminated during the running stage. Another characteristic of the response/selection planning subsystem is that, unlike the sensory analysis and response execution subsystems which continually process new information and update their outputs, the response selection/planning subsystem only updates its output intermittently as indicated above. That is, the response selection/planning subsystem only updates its output when it has finished processing the information that must be processed in order to determine the next lowest level response.

The final subsystem is the response execution subsystem which takes the motor plan for the lowest level subresponse and translates it into the many signals that activate the appropriate muscle groups, which in turn provide the appropriate movement. To generate the signals for the muscles, the response execution subsystem must first take from the output of the response selection/planning subsystem the plan which describes the required subresponse. This subresponse plan describes the movement in abstract terms as a function of time. For example, the plan may be to accelerate the forearm at a particular acceleration for a specific time, to then move the arm at a constant velocity for a predefined time, and then to deaccelerate the arm for a time sufficient to bring it to a stop. The response execution subsystem then applies this subresponse plan to an inverse model of the musculoskeletal system which has the desired

response as its input and the many muscle signal channels as its output (Neilson *et al.*, 1992). Thus, the inverse model produces a set of continuous signals, from the subresponse plan, that activate the appropriate muscles at the appropriate time to produce the required response.

The inverse model of the musculoskeletal system is an adaptive model as it changes its parameters as the loading on the musculoskeletal system changes and as the state of the muscles and joints change with various movements. The current state of the muscles and position of the joints is conveyed continually to the model via proprioceptive feedback. Proprioceptive feedback is the process by which the human body perceives its position, and that of its components, in space and how this position changes with movement. As the model is adaptive it is possible with repeated exposure to a particular task to refine the model parameters that describe the musculoskeletal system in the particular situation and to become very skilled at performing the task. It is also noteworthy that the response execution subsystem is considered to be an open-loop system as once the system has received the required subresponse plan the movement defined by the subresponse plan cannot be stopped or changed during its execution (Neilson *et al.*, 1992).

The overall performance of the S-M system in performing a particular task is clearly dependent on the integrity of all the systems described above which together form the S-M system. A person's concentration, motivation, intelligence, and previous exposure to a task can, however, also have a marked influence on S-M performance. In the situation where the integrity of the S-M system is seriously and permanently compromised, one of the common causes is a neurological disorder. Neurological disorders are those disorders which impair the function of the nervous system and, as the nervous system is a major component of the S-M system, the S-M system is often impaired. As the integrity of the S-M system is very important if a person is to function effectively in everyday life, the investigation of these neurological disorders is of great interest. Investigation of abnormal S-M function allows treatment procedures to be developed and assessed, the progress of disorders to be tracked, and the quality of life of an affected person to be improved. The major involvement of the CNS in the correct functioning of the S-M system is also fortuitous in that it provides a window to the brain. By assessing both normal and abnormal S-M function it is possible to assess how the brain functions in both its normal and impaired states.

8.2 PARKINSON'S DISEASE

There are many neurological disorders that can seriously impair the functioning of the S-M system. In this section Parkinson's disease, a neurological disorder which can seriously affect the S-M system, is described. This particular disorder is described here because subjects with this disorder were used, during the author's research, to determine whether the 2-D tasks developed could correctly demonstrate the deficits known to occur in the Parkinsonian S-M system and to determine whether 2-D tasks could provide any new information on the effects of the disease. The description begins with a brief overview of the pathophysiology of the disease, followed by a reasonably detailed look at the symptoms of the disease which focuses on those symptoms which seriously degrade the performance of the S-M system. Finally, a brief account of how the disease can be treated is given.

Located at the base of each hemisphere of the forebrain is a gray mass known as the basal ganglia. The basal ganglia comprises three substructures known as the caudate nucleus, globus pallidus, and the putamen and is known to be involved in the control of motor function (Carlson, 1986). In the midbrain there is a dark staining structure known as the substantia nigra.

The axons of some of the neurons in the substantia nigra terminate in the caudate nucleus and these axons form a neural pathway known as the nigrostriatal bundle (Carlson, 1986). The rate at which neurons fire in the brain is controlled by chemical substances known as neurotransmitters. In the basal ganglia there are primary two neurotransmitters, dopamine and acetylcholine which are inhibitory and excitatory neurotransmitters respectively (Carlson, 1986). Dopamine is produced in the dopaminergic neurons that form the nigrostriatal bundle. If a significant portion of the cells which form the nigrostriatal bundle die or there is a significant reduction in the dopamine release capacity of the dopaminergic neurons, a dopamine/acetylcholine imbalance occurs in the basal ganglia. This imbalance in the concentration of the neurotransmitters is known to cause the symptoms which characterise Parkinson's disease. That is, Parkinson's disease is caused by a dopamine/acetylcholine imbalance in the basal ganglia.

The symptoms of Parkinson's disease which disrupt the performance of the motor system are (Eyzaguirre and Fidone, 1975; Carlson, 1986):

- Bradykinesia, which is the slowing of movement. People with Parkinson's disease find it takes them much longer than normal to perform activities required in everyday living.
- Akinesia, which is a resistance to begin movement and when moving a resistance to stop moving (i.e. the motor system has a large amount of inertia). People with Parkinson's disease find it difficult to carry out tasks that require rapid changes in movement, such as pacing backwards and forwards across a room (Carlson, 1986).
- Hypokinesia, which is a reduction in the range of movement. This symptom can be seen in Parkinsonian handwriting which becomes smaller and less well formed as writing progresses.
- Muscular rigidity, which is the continuous contraction of muscles within the body. Muscular rigidity has the effect of making the joints appear stiff and in some cases gives people experiencing Parkinson's disease a mask-like appearance.
- Tremor, which is the involuntary shaking of the limbs. This tremor has a frequency of 5-8 Hz and is usually most pronounced when the person is not performing any movement but is attentive and alert. When the person is relaxed and unoccupied the tremor tends to reduce and during sleep it disappears completely. As the disease progresses some people develop a noticeable tremor of 8-12 Hz that occurring during voluntary movement.

The above symptoms tend to be unilateral initially but as the disease progresses they usually become bilateral. Other characteristics of Parkinson's disease are abnormalities in posture, a loss of certain automatic associative movements (e.g. swinging arms when walking), muscular deformities, and speech difficulties (Eyzaguirre and Fidone, 1975).

While Parkinson's disease was originally believed to be a disorder that primarily affected the function of the motor system, there is a growing amount of evidence that indicating the disorder also affects the sensory analysis and cognitive functions of the brain. The deficiencies in cognitive function include a reduced ability to switch between two different non-motor tasks, a reduced ability to update a response strategy based on the effect a response has had in the past (e.g. performance on the Wisconsin card sort test is degraded (Schmidt, 1988)), and an impairment in associative learning tasks (Schmidt, 1988). With the sensory analysis system two main deficits are observed. These are the intrusion of extraneous sensory information into a motor plan (i.e. any extraneous sensory stimuli can halt an ongoing movement) and a reduced ability to incorporate proprioceptive feedback into motor plans (e.g. some people

with Parkinson's disease are unaware of postural disturbances or are unable to determine when a dental impression of their own teeth is correctly aligned in their mouth (Schneider and Lidsky, 1987)). In addition, performance of the visual system is also degraded in Parkinson's disease. A reduction in visual acuity and contrast sensitivity are just two of a number of deficits that have been observed in Parkinson's disease (Jones *et al.*, 1992). The above discussion indicates that the subsystems predominately affected in Parkinson's disease are the response execution and sensory analysis subsystems (Fig. 8.1), although the effects are not limited to these areas.

There are a number of pharmacological agents which can be used to correct the acetylcholine/dopamine balance in the basal ganglia and to reduce the effects of the symptoms. There are three basic treatment strategies and these are: to boost the dopamine concentration in the basal ganglia by administering the precursor to dopamine (i.e. levdopa), to decrease the sensitivity of the neurons in the basal ganglia to acetylcholine by using anticholinergic drugs, and to increase the sensitivity of the neurons to dopamine by using dopamine agonists. In addition other pharmacological agents are also prescribed to reduce the effects of the symptoms, to reduce side effects of the above agents, and to improve the mental state of the person with Parkinson's disease.

8.3 MEASUREMENT OF SENSORY-MOTOR FUNCTION

In many clinical and research fields it is necessary to be able to measure and assess the performance of the S-M system and its components. In the clinical arena the need arises from the necessity to be able to diagnose disorders which adversely affect the S-M system and to assess the success of various treatment strategies. In the area of research the motivation for wanting to measure S-M function differs with each field of research. Common areas of S-M research and their motivation include: medical research where the motivation is to gain an understanding of the physical causes and effects of various S-M disorders and to thereby devise treatments; psychomotor research where there is a desire to gain an understanding of how the S-M system works in both its normal and abnormal states; human performance measurement where there is a desire to measure the limits and characteristics of human motor performance and, in some cases, devise methods to extend performance boundaries; and in the human operator area where the motivation is to be able to develop models of the S-M system. The above research areas are, however, not defined by strict boundaries and it is common for S-M research to cover several of the above fields.

In all of the above research areas numerous tests have been developed to allow performance measurements to be made on various S-M system components and to thereby permit an understanding of various aspects of the system to be gained. The type of test developed and used in each area depends on the type of information required. For example, different tasks are usually required to measure either component or integrated S-M function, to measure the performance of different S-M system components, to provide either a qualitative or quantitative assessment of function, and to provide either objective or subjective assessments of performance. For most scientific research on the functioning or dysfunctioning of the S-M system, tests capable of providing objective and quantitative measures of S-M performance are generally required. In this section an overview of some of the tests used to measure S-M function is given. The overview begins with a brief account of the classical neurological tests, which provide subjective and qualitative measures of performance, and moves on to quantitative tasks, such as tracking tasks, which provide more refined and detailed measures of S-M performance.

8.3.1 Conventional Clinical Tests

Clinical tests routinely used by medical professionals to evaluate the functioning of the S-M system are generally subjective in nature. In this situation, the primary purpose of the S-M evaluation is to detect and diagnose the presence of a disorder or to assess the success of treatment and a skilled judgement on the part of the medical professional is sufficient. In Table 8.1 some common tests used by neurologists, physicians, nurses, and occupational therapists to assess the integrity of the S-M system are listed to give an indication of the type of S-M assessment usually performed clinically. This is by no means a complete list of the tests used clinically to assess the S-M system and the interested reader is referred to Rosenthal *et al.* (1983), Perkin (1986), Bearce and Myers (1990). From Table 8.1 it can be seen that the majority of the tests require the examiner to make a subjective judgement on whether the S-M system is functioning correctly. In the situation where the tests indicate that the S-M function is abnormal, the examiner must also make a subjective judgement of the degree of abnormality.

The subjective nature of most of the clinical S-M tests leads to a lack of sensitivity, intra- and inter-examiner inconsistencies, unreliability, and limited use of the results in scientific research. Medical professionals recognised this weakness with many of their clinical tests and, to overcome this weakness, developed ordinal classification schemes (Kondraske *et al.*, 1984). In an ordinal classification scheme the continuum of S-M performance on a particular task is divided into a set number of ranked categories. Each of the categories has associated with it a description which characterises the properties that the observed performance must possess to be included within that category and a number (grade) which describes the position of this category in the rank of categories. An example of an ordinal classification scheme is given in Table 8.2 where the scheme used to classify tendon reflexes is given (Perkin, 1986; Bearce and Myers, 1990). Examples of other ordinal classification schemes are the MRC scale for assessing motor power (Perkin, 1986), the Glasgow Coma scale for assessing states of consciousness (Rosenthal *et al.*, 1983; Perkin, 1986; Bearce and Myers, 1990), the Hoehn-Yahr scheme for classifying the severity of Parkinson's disease (Hoehn and Yahr, 1967), and the Northwick Park ADL scale for assessing activities of daily living (Sheikh *et al.*, 1979).

Ordinal classification schemes allow medical professionals to obtain a more objective assessment of the performance of the S-M system. As the different quality classifications are linked to numerical values it is now possible to obtain an overall assessment of S-M performance by summing the grades obtained from various ordinal classification schemes (see the comprehensive S-M ordinal classification scheme in Rosenthal *et al.* (1983)). It is also possible to apply statistical techniques to the results obtained from different groups of subjects to determine whether any differences observed in performance are significant. There is still the weakness, however, that the assessment relies on the examiner making a subjective evaluation of performance, although this evaluation is now done within the bounds of some guidelines. Thus, ordinal classification can reduce unreliability and inter- and intra- examiner variability but cannot eliminate these undesirable factors completely. There are also a number of other limitations of ordinal classification schemes (Kondraske *et al.*, 1984):

- The accuracy of the ordinal classification scheme is dependent on the ease with which the various categories of performance can be described by a definitive statement of characteristics.
- The continuum of S-M performance is usually divided up into a small number of rather broad categories (i.e. usually less than six) and for this reason these classification schemes cannot detect subtle differences in performance.

CLINICAL TESTS OF SENSORY-MOTOR SYSTEM INTEGRITY

Function Tested	Test Type	Test/Description
Sensory Evaluation		
Visual System.	Qt.	Snellen chart for visual acuity.
	Ql.	Response of pupil to light.
	Ql.	Colour vision.
	Ql.	Visual field by confrontation.
Auditory System.	S.	Watch ticking test and whispering test for a gross estimate.
	Ql.	Weber and Rinné tests: assess hearing using tuning forks placed about head.
Olfactory System.	Ql.	Identification of aromatic substances with eyes closed.
Tactile System.	Ql.	Distinction of light touch, dull pressure, and sharp pain.
	Ql.	Identification of common objects held in hand with eyes closed.
Proprioceptive System.	Ql.	Identification of direction of movement around a joint, with eyes closed. Standing on two feet or one foot (eyes closed).
Motor Evaluation		
Reflexes.	S.	Percussion of appropriate tendon.
Motor Strength (isometric).	S.	Ability to push against resistance provided by examiner. Usually test major muscle groups.
Muscle Tone.	S.	Resistance to passive movement of a limb.
Range of Movement.	Qt.	Measure joint range of movement using a goniometer.
Integrated Motor Function Evaluation		
Coordination.	S.	Rhythmic point to point tasks such as sequentially touching finger on each hand with thumb on that hand (simultaneously on both hands, backwards, and forwards), and alternately tapping top and bottom of hands on thighs.
	S.	Finger to nose test.
	S.	Run heel from knee to shin on opposite leg.
	S.	Tap hand of examiner with toes as rapidly as possible.
	S.	Multiple finger 'drumming'.
Activities of Daily Living.	S.	Assess ability to perform activities of daily living such as self care and home skills.
Other Procedures		
Higher Mental Function.	Ql.	Memory Tests.
	S.	Free-Hand drawing.
	S.	Copying.
	Ql.	Orientation in place and time.
External Physical Assessment.	S.	Evidence of tremor.
	S.	Gait/posture abnormalities.
Internal Physical examination.	S/Qt.	NMRI, CT, EEG, EMG see Table 1.1.
	Qt	Lumbar puncture.

Table 8.1 Key: S. = Subjective, Ql. = Qualitative, and Qt. = Quantitative. Subjective tests are those where the examiner assesses the degree of abnormality without using a measuring device (i.e. the assessment is based solely on the examiner's judgement). Qualitative tests are those where the test unequivocally indicates that a function is abnormal or normal but provides no measurement of the degree of abnormality. With quantitative tests a measuring device is used to determine the degree of abnormality.

ORDINAL CLASSIFICATION SCHEME FOR GRADING OF TENDON REFLEXES

Grade	Classification description
5	Reflex pathologically brisk
4	Reflex brisk but still within normal limits
3	Reflex normal
2	Reflex just present
1	Reflex present only with reinforcement
0	Reflex absent

Table 8.2

- While the schemes can be used in conjunction with statistical techniques to detect difference in performance, the magnitudes of any difference cannot be determined because the numerical grades are a measure of rank and not a physical quantity.
- Classification schemes only provide information on how well a task is performed and no information on how the task was performed.

To allow the performance of the S-M system to be evaluated in a completely objective manner and to overcome most of the above limitations associated with ordinal classification schemes, researchers devised tests capable of providing a quantitative evaluation of S-M performance.

8.3.2 Quantitative Tests

Psychologists and other researchers recognised the need for tests that could provide quantitative measures of S-M performance. Quantitative tests can usually provide reasonably sensitive measures of physical S-M quantities and can be used in conjunction with statistical techniques to provide an indication as to the size and nature of any differences observed in the S-M function of different groups. In Table 8.3 a series of non-computerised tests that can be used to provide quantitative measures of various S-M functions are tabulated. This table does not provide a complete list of tests used to measure various aspects of S-M function but provides a reasonable overview of the type of tests used. More complete lists and descriptions of quantitative S-M function tests can be found in the work of Fleishman (1972), Potvin and Tourtellotte (1975), and Kondraske *et al.* (1984). Of special note is the work of Fleishman (1972) (Fleishman, 1954; Fleishman, 1958; Fleishman and Ellison, 1962) who investigated 200 different tests to determine what fundamental S-M functions could account for the performances observed with this wide range of tasks. From these studies, Fleishman (1972) found that the eleven psychomotor functions listed in Table 8.3, in conjunction with nine physical proficiency abilities (i.e. dynamic strength, static strength, explosive strength, trunk strength, dynamic flexibility, static flexibility, gross body coordination, gross body equilibrium, and stamina), could consistently account for the performance observed with the tests. It was also found that those tests listed in Table 8.3, in addition to some others, provided the best measure of the indicated S-M function, e.g. control precision was found to be most effectively measured by the rotary pursuit task. The most important attribute to note about the tests listed in Table 8.3, and those other tests that research psychologists regularly use listed in Fleishman (1972), (Fleishman, 1954; Fleishman, 1958; Fleishman and Ellison, 1962), is that they all provide quantitative measures of S-M performance. Thus, statistical techniques can be used in conjunction with the results

obtained from these tests to obtain conclusive and quantitative results on the function of the S-M system and its subsystem in various states.

COMMON QUANTITATIVE S-M TESTS

Function Tested	Test	Description	Quantitative Measure
Control Precision.	Rotary Pursuit.	A rotating metallic disc embedded in a turntable must be followed with a stylus held in the hand.	Time-on-target.
Multilimb Coordination.	Two-Hand Pursuit.	A metallic target moves in an irregular manner on a rotating disk. The target position must be matched with a stylus controlled independently in orthogonal directions by each hand.	Time-on-target.
Response Selection.	Choice Reaction Time.	A set of lights are spatially located in a pattern that corresponds to a set of switches. When a light is turned on the corresponding switch must be pushed as quickly as possible.	Cumulated Reaction times.
Reaction time.	Visual/audio reaction timer.	When a single light or buzzer is turned on a switch must be pressed as quickly as possible.	Reaction time.
Speed of arm movement.	Two plate tapping.	Two adjacent metal plates, separated by at least 300 mm, must be alternatively tapped with a sylus as quickly as possible.	Taps/time.
Rate Control.	Motor Judgement.	The speed of a rotating pointer situated between two rotating disks, with alternate white and black sectors, is controlled so that the pointer revolves as many times as possible in the completely white sectors.	Revolutions/time.
Manual Dexterity (i.e. large movements).	Minnesota range of Movement.	Large blocks must be removed from holes, turned over with the other hand, and then placed back in the holes.	Number/time.
Finger Dexterity (i.e. small movements).	Purdue Pegboard.	Put small pegs in a series of holes as rapidly as possible.	Number/time.
Arm-Hand Steadiness.	Track tracking.	A stylus encircling a wire, which is bent into an irregular shape, must be moved along the wire without touching the wire	Number of touches.
Wrist finger speed.	Tapping.	A pencil is used to make three dots in as many 10 mm diameter circles as time permits	Number of completed circles.
Aiming.	Aiming.	A pencil is used to make a dot in as many small (3mm diameter) circles as time permits	Number of correct dots.

Table 8.3

One of the major problems with the quantitative tests in Table 8.3, and those reported in Fleishman (1972) routinely used by psychologists, is that they test the complete S-M system or a substantial portion of the system. A major area of research in neurology is concerned with the dysfunction in fundamental S-M subsystems and for this reason neurological researchers require tests that can assess both component and integrated S-M function. To design test batteries that can examine the gamut of S-M subsystems, the S-M system can be divided up conceptually into its fundamental subsystems and tests devised to test each fundamental subsystem in isolation. To divide the S-M system up into its subsystems it is common to take a particular integrated S-M task and then hypothesise what S-M subsystems must be used to perform the task. For example, performance on a visual reaction time task is dependent on the integrity of the visual system, sensory analysis subsystem, response planning subsystem, proprioceptive feedback, upper limb speed, etc. It is impossible to separate all S-M subsystems into their fundamental components, particularly within the brain, and so the aim is to obtain as much S-M independence as possible. One of the most comprehensive test batteries designed to measure both component and integrated S-M function in the neurological context is that designed by Potvin and Tourtellotte (1975). The test battery was conceived in the 1960s and

was developed to the stage where it contained fifty-seven instrumented tests of component and integrated S-M function. The S-M test battery was designed to test S-M function as it relates to the upper and lower limbs, head and neck, and body (Kondraske *et al.*, 1984). A sample of the tests included within this battery are those to measure mental function, vision, hearing, touch, strength, coordination, reactions, tremor, and steadiness. A more detailed description of the test and instrumentation can be found in Potvin and Tourtellotte (1975) and Kondraske *et al.* (1984). In the early 1980s this neurological test battery was integrated and automated through computerisation (Kondraske *et al.*, 1984). The test battery has been used extensively to investigate aspects of Huntington's disease, the effects of hypnosis, the effects of learning, and the effectiveness of various neuropharmological compounds used to treat Parkinson's disease (Kondraske *et al.*, 1984).

Another limiting aspect of the traditional tests used by psychologists to measure S-M function is that they provide very little information about the S-M strategy used to complete the test and use non-ideal measures of performance. This is particularly true of tests used to measure integrated S-M function, such as the pursuit rotor and two-hand pursuit task. These tests measure the integrity of the S-M system by using measures such as time-on-target. This has, however, been shown by Poulton (1974) to be non-ideal for measuring S-M function on the basis that it, and many of the other measures in Table 8.3, do not provide a measure of the error between the ideal response and the actual response which is continuous in space. With the time-on-target error measure in the pursuit task, a subject who is completely unable to follow the target but makes contact a number of times by chance is considered, by this measure, to perform better than a subject who follows the target movement correctly but is always just outside the target disk. Also a subject whose response is always just inside the target disk is considered to perform much better than a subject whose response is just outside the disk. The above scenario can occur with other tests in Table 8.3 such as the aiming and tapping tests. The lack of a continuous error measure also means that it is impossible to determine quantitatively what strategy was used to complete any of the tests in Table 8.3.

The developments in modern computers and electronics have now made it possible, however, to in effect continuously record the movements made by a subject when completing an integrated S-M task. By using these recordings of the movements it is possible to derive more suitable measures of performance and, by using various signal analysis techniques, it is possible to assess various aspects of the response, e.g. the strategy used to complete the task. A group of tests that have benefited most from these developments are the group of tests known as tracking tasks. Numerous experimental studies using tracking tasks have shown that these tasks combined with suitable performance analysis techniques can provide valuable quantitative information on the S-M system and its subsystems. It is worth noting that the neurological test battery of Kondraske *et al.* (1984), described earlier, includes a series of computerised tracking tasks of this type. In the following section tracking tasks are described in terms of their structure, history, applicability, usefulness, and type of information they can provide on S-M function through appropriate signal analysis. The reason for describing tracking tasks in detail is that the development of two tracking tasks is a significant component of the original work contained within this part of the thesis.

8.3.3 Tracking Tasks

Tracking tasks have been used extensively since the second world war to provide quantitative information on the performance of the S-M system and to determine how the system functions in both its normal and diseased states (Schmidt, 1988). Tracking tasks provide this quantitative

assessment of S-M function by assessing the ability of a subject to follow a moving visual target with a pointer controlled by coordinated movement of a limb. The major constituents of a tracking task are a device used to produce the target signal, a device used to display both the target signal and the response by the subject (e.g. a oscilloscope or computer monitor), a device used to measure the subject's motor response (e.g. steering wheel, joystick, or mouse), and a device to compute the performance measures. The important attributes which make tracking tasks so valuable for the study of S-M function are that they are objective and quantitative, can provide detailed information on how the task was performed, make continuous demands on many S-M subsystems, can provide information on the integrity of many S-M subsystems (on applying appropriate analysis techniques), are similar to activities performed in everyday life (e.g. driving), and allow considerable control and flexibility over the target signal and the motor output device type and characteristics.

In this section a brief review of tracking tasks, and their characteristics and areas of application is given. This review should provide the reader with an appreciation of the value of tracking tasks in providing information on both normal and abnormal S-M function. This review is primarily limited to computerised hand/eye 1-D pursuit tracking tasks, although non-computerised tasks are discussed in the historical review of tracking tasks. In a pursuit tracking task, the target signal is displayed to the subject and the subject must follow the movement of the target signal as closely as possible with the pointer which they control (see Fig. 8.2(a)). This is in contrast to compensatory tracking tasks, where the error signal is displayed to the subject and the subject must then, through the coordinated movement of a motor output device, attempt to reduce the displayed error to zero (see Fig. 8.2(b)). Two-dimensional pursuit tracking tasks (tasks where the target signal and response have two spatial degrees of freedom) are described in the following section as the tasks developed by the author are of this type and a separate treatment of these tasks is therefore considered appropriate.

An interest in how the human operator functions as part of a closed-loop control system led to the development of the first pursuit tracking tasks around the time of the second world war. The primary concern of researchers at this time was to improve the performance of aircraft and weapons control system by gaining an understanding of the characteristics of the human component within these systems (Craik, 1948; Helson, 1949; Warren *et al.*, 1952). The tasks of this era were electromechanical devices, with performance analysis being performed manually or by simple electrical circuits (e.g. to measure time-on-target, a clock begins timing when electrical continuity is established between the target and the subject's stylus). The rotary pursuit task and two-hand pursuit task described in Table 8.3 were tasks developed in these early days and, like these two tracking tasks, many of the other electromechanical tasks used non-ideal performance measures (e.g. time-on-target see §8.3.2). One of the first tracking tasks to use more appropriate performance measures was developed by Warren *et al.* (1952). It was was based around an oscilloscope with electronic performance analysis circuits which, in addition to measuring time-on-target, could provide measures of the mean error and mean square error. These performance measures are explained later in this section. During this period there was also considerable interest in using tracking tasks to investigate aspects of the psychomotor system such as learning (Adams, 1952; Slack, 1953; Archer *et al.*, 1956), motor planning (Poulton, 1952), and the ability to perform multiple tasks (Jackson, 1958; Garvey and Taylor, 1959).

The advent of powerful and versatile modern computer systems has provided an ideal platform for tracking tasks due to their ability to generate real-time displays, store large amounts of raw data, and carry out fast analysis and feedback of performance. S-M system analysis is one area that has benefited enormously from the processing capabilities of computers,

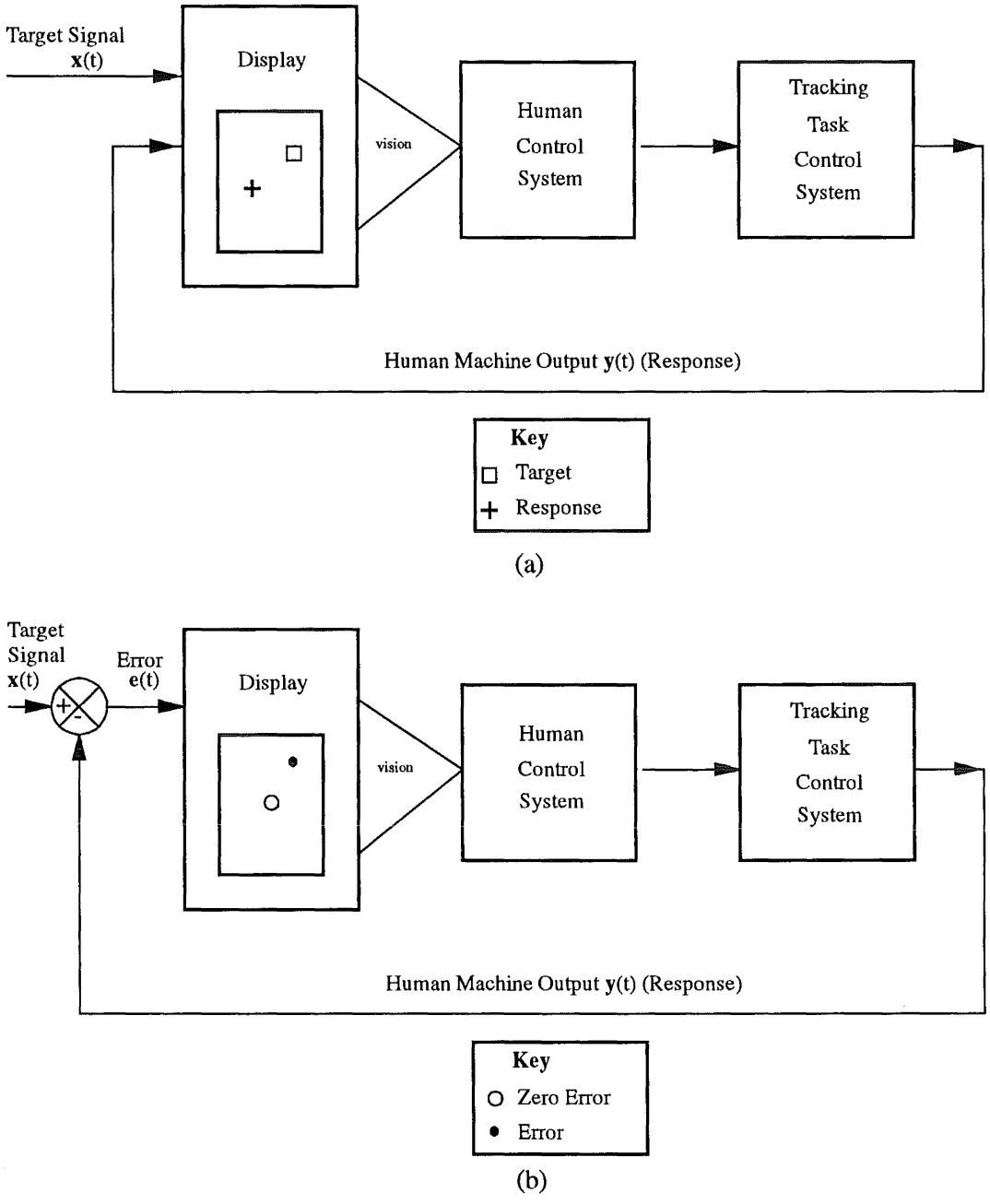

Figure 8.2 (a) Schematic of a pursuit tracking task where the subject aims to reduce the error between the displayed target and response signals, i.e. minimise (x(t) - y(t)). (b) Schematic of a compensatory tracking task where the subject aims to reduce the displayed error to zero, i.e. minimise e(t).

since more appropriate measures of performance can be computed and, by applying various signal analysis techniques, it is now possible to analyse further the function of the S-M system and its subsystems. The modern computer has enabled researchers to integrate tracking tasks and other psychomotor tests into one unit, forming comprehensive batteries for the evaluation of the S-M system and its subsystems (Kondraske *et al.*, 1984; Jones, 1986b; Jones *et al.*, 1989). The quantitative nature and versatility of computerized tasks has led to their extensive use in medical and rehabilitation research for the study of recovery following stroke (Corkin, 1968; Lynn *et al.*, 1977; De Souza *et al.*, 1980; Jones and Donaldson, 1981; Jones *et al.*, 1989; Jones *et al.*, 1990), the effectiveness and side effects of drugs (Cassell *et al.*, 1973; Pullman *et al.*, 1988; Jones *et al.*, 1989), Parkinson's disease (Cassell *et al.*, 1973; Bloxham *et al.*, 1984; Frith *et al.*, 1986; Abdel Malek *et al.*, 1988; Behbehani *et al.*, 1988; Pullman *et al.*, 1988; Jones and Donaldson, 1989), cerebral palsy (Riley and Rosen, 1987; Nash *et al.*, 1989; Neilson *et al.*, 1990), multiple sclerosis (Riley and Rosen, 1987), cerebellar ataxia (Riley and Rosen, 1987), Alzhiemer's disease (Baddeley *et al.*, 1986), head injury (Jones and Donaldson, 1981; Riley and Rosen, 1987), as well as their application in the development and testing of devices for the disabled (Riley and Rosen, 1987; Ison, 1988), and in driving assessment (Jones *et al.*, 1983; Croft and Jones, 1987). Researchers are also using computerized tasks to develop models of the normal and impaired S-M system (Lynn *et al.*, 1979; Abdel Malek *et al.*, 1988; Neilson *et al.*, 1988; Neilson *et al.*, 1992).

Modern computerised tracking tasks use many measures of performance and those performance measures which are most commonly used to analyse S-M function are now described. The strengths and weaknesses of the various performance measures are also described along with the information that they can provide on the S-M system and its function. The reader who is particularly interested in a detailed account of the performance measures that can be used in conjunction with tracking tasks are referred to Poulton (1974). When deciding on a performance measure it is important to remember that, if the purpose of the tracking task is to compare the S-M performance of different groups, the vast amount of information obtained from tracking tasks must be reducible to a manageable from and the performance measures must be suitable for statistical analysis.

The most useful and popular measures of tracking performance are those based around the error. The error is the signed difference between the target signal and the subject's response pointer at the time when the subject's response is sampled. Plots of the error as a function of time can provide information on the learning process, intra-task changes in the subject's level of concentration and motivation, and those sections of the target signal which the subject finds most difficult to track. To reduce the errors to a manageable form it is usual to compute some form of average of the errors as this provides an overall measure of performance. The simple mean is, however, of limited use and in effect only indicates whether more errors are made on one side of the target than the other. The most commonly used average error measures are the mean absolute error (MAE) and the root mean square (RMS) error where the signs of the errors are in effect removed. These two measures provide a useful assessment of performance as any variability in the subject's response is not cancelled out and comes through in the performance measure (Poulton, 1974). Unfortunately the mean error, which for a group of subjects tends to have a Gaussian or Normal distribution, is most compatible with standard parametric statistical techniques (Poulton, 1974). The distributions from the other error measures cannot be Gaussian and tend to be skewed since it is now impossible to have a negative mean error and it is very unlikely that any subject could obtain a MAE of zero. Thus, nonparametric statistical techniques are often considered more appropriate for analysis of results derived from performance measures such as the MAE and RMS error,

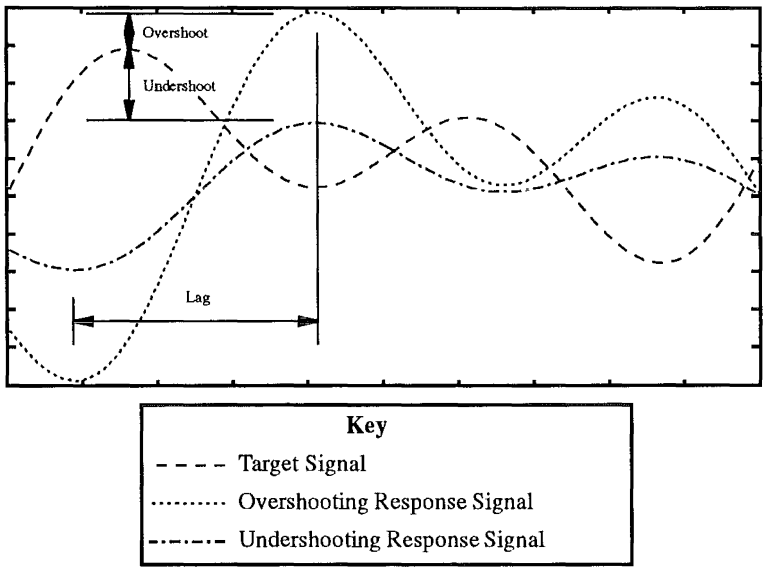

Figure 8.3 A tracking task target signal and two different types of possible response signal. The two responses shown demonstrate a response which lags and overshoots the target signal and a response which lags and undershoots the target signal. The gain provides an estimate of this global over- or under- shoot and the mean time delay provides the lag.

Important information on the tracking strategy used by the subject can be obtained by computing the ratio of the RMS response signal amplitude to the RMS target signal amplitude. The value obtained is known as the gain and indicates whether the subject tends to either overshoot or undershoot the target (see Fig. 8.3). If the subject tends to undershoot the target they may be *cutting corners* because they find it difficult to keep up with the target or cannot provide the dynamic range of movement required. The final time-domain performance measure which is commonly used is time-on-target. For the reasons outlined previously in §8.3.2 the time-on-target measure is considered to be an unsuitable performance measure.

On analysing the subject's response in the frequency domain, it is possible to derive performance measures such as the magnitude and phase response of the subject's S-M system. To obtain the magnitude response, the ratio of the RMS magnitudes of the frequency components comprising the subject's response signal and the corresponding frequency components in the target signal are plotted as a function of frequency. Plots of the magnitude response of the S-M system are of interest when developing mathematical models of the S-M system and when attempting to determine the bandwidth of various S-M subsystems. The gain of the subject's response can also be determined from the S-M system's frequency response. Plots of the phase difference between the target and response signals as a function of frequency also provide important information on the strategies employed by the S-M system to complete a task. If the response signal tends to lead the target signal, the subject was able to predict the future position of the target and plan a suitable anticipatory response. With random target signals the subject's response usually lags behind the target. There are a number of inherent delays within the S-M system that contribute to these observed phase differences. These are physical delays in the system (e.g. the time taken for a signal to propagate along a nerve) and processing delays in the sensory analysis subsystem, response selection/planning subsystem, and response execution subsystem. To gain an estimate of the overall delay in the S-M system it is usual to use cross-correlation techniques to obtain an estimate of the mean time delay (i.e. the time delays from each frequency component are averaged). Poulton (1974) suggests,

however, that by averaging the time delays information on the tracking strategy is lost and for this reason Poulton does not recommend this performance measure. Nonetheless, the average time delay is a valuable measure of overall performance and is suitable for statistical analysis. It has also been noted by Cassell (1973) that although the subject's delay may change during the course of the task the changes have a negligible effect on the width of the peak of the cross-correlation function, which defines the time delays, and the time delays observed in each frequency component tend to be constant.

8.3.4 2-D Tracking Tasks

Two-dimensional tracking tasks are a logical extension to 1-D tracking tasks. With a 2-D tracking task the subject must follow the movement of a target moving in 2-D space with a pointer controlled from a single input device. In contrast, Tracking tasks with two 1-D targets, moving in orthogonal directions, which must be followed separately by coordinating the movement of either two limbs or a single limb are called dual axis tasks (see Navon *et al.* (1984) and Fracker and Wickens (1989)).

The author's interest in 2-D tracking tasks stemmed from a belief that 2-D tasks may stress the S-M system to a greater extent than 1-D tasks and therefore maybe able to provide either new, additional, or more sensitive information on S-M function. In addition, the multidimensional nature of 2-D tasks makes them more closely allied to the multidimensional environment in which the S-M system must function. Thus, it is believed that 2-D tasks may provide more valid information on S-M function than 1-D tasks. While 2-D tasks have the potential to provide additional information on S-M function, they have not been as widely used in S-M research. The reason for this could be that, to the author's knowledge, no-one has attempted to determine whether there is any benefit to be gained from using 2-D tasks in preference to 1-D tasks. Furthermore, in the time before computerisation of tracking tasks became viable, the multidimensional nature of 2-D tracking tasks made performance analysis difficult.

The original research contained within this part of the thesis is concerned with determining whether the beliefs outlined above are true and whether there is any benefit to be gained from using 2-D tasks in preference to 1-D tasks. In this section a brief review of the areas of research where 2-D pursuit tracking tasks have been used in the past is given and the general form of the tasks is described. The performance analysis techniques that can be used in conjunction with 2-D tasks are then described and following this the research of others which indirectly may help to either substantiate or refute the above beliefs is reviewed.

Advances in computer technology over recent years has made it considerably easier to generate 2-D tasks and to analyse the results. This has lead to a considerable increase in the number of 2-D computerised pursuit tracking tasks used for for general S-M research (Kondraske *et al.*, 1984; Trier Andersen, 1986), psychomotor research (Frith and Lang, 1979; Williges and Williges, 1978; Viviani, 1988; Viviani and Mounoud, 1990), S-M system modelling (Viviani *et al.*, 1987; Viviani and Mounoud, 1990), and medical research (Bloxham *et al.*, 1984; Baddeley *et al.*, 1986; Frith *et al.*, 1986; Cooper *et al.*, 1989). The majority of the tests used by the above researchers required the subject to track either a square or circular target moving on either a oscilloscope or computer monitor through movement of an upper-limb. The target movement was defined by either a random signal (Williges and Williges, 1978; Kondraske *et al.*, 1984; Viviani *et al.*, 1987; Trier Andersen, 1986; Baddeley *et al.*, 1986) or by a summation of three sinusoids (Bloxham *et al.*, 1984; Frith *et al.*, 1986). The subject's response was displayed on the screen as either a cross or a dot and the transducer used to measure the subject's motor

output was a joystick, except Baddeley *et al.* (1986), who used a light pen, and Viviani *et al.* (1987). Performance analysis was computerised in all cases. The 2-D task used by Viviani *et al.* (1987) was novel in that a computer was used to control the position of a laser spot on a transparent digitising table. The subject then had to follow the movement of this laser spot on the digitising table with the digitising table's puck.

The performance measures usually used in conjunction with 2-D tracking tasks are the same as those described in §8.3.3 for 1-D tasks. Typically the 2-D target and response signals are separated into their vertical and horizontal components (x and y coordinates) and the 1-D performance measures are applied to the resulting 1-D signals. All of the 2-D tracking tasks outlined above, except for that of Viviani *et al.* (1987), used various combinations of the error, MAE, gain, mean time delay, and phase difference at various frequencies to measure the tracking performance. A few researchers also used the time-on-target measure. In contrast Viviani *et al.* (1987) used the equally valid approach of using vector quantities to describe performance. Viviani *et al.* (1987) notes that representing the performance vectorially (as opposed to orthogonal components) may be the most relevant representation, as the S-M system is likely to code the visual error and motor response in vectorial terms. The vectorial performance measures used were the position error vector $\mathbf{e}(t)$ and the difference, $\mathbf{v}(t)$, between the tangential target and response velocity vectors. If the 2-D target signal is denoted by $\mathbf{x}(t)$ and the 2-D response signal is denoted by $\mathbf{y}(t)$, $\mathbf{e}(t)$ and $\mathbf{v}(t)$ are given by

$$\mathbf{e}(t) = \mathbf{x}(t) - \mathbf{y}(t)$$

and

$$\mathbf{v}(t) = \frac{d\mathbf{x}(t)}{dt} - \frac{d\mathbf{y}(t)}{dt}$$

respectively. To obtain an overall measure of performance from the vectorial performance measures the means of the magnitudes of these vectors could be computed. In addition to the above error measures, Viviani *et al.* (1987) computed an instantaneous delay $d(t)$ defined to be the value $d(t)$ that minimises the expression

$$E^2(t) = \frac{1}{2\epsilon}\int_{-\epsilon}^{\epsilon} |\mathbf{x}(t+\tau) - \mathbf{y}(t - d(t) + \tau)|\, d\tau,$$

where ϵ is a preset interval over which the integration is performed. Once $d(t)$ has been calculated, the delayed velocity difference

$$\mathbf{v}_d(t) = \frac{d\mathbf{x}(t)}{dt} - \frac{d\mathbf{y}(t - d(t))}{dt}$$

can be computed, along with another overall measure of performance known as the global distortion,

$$\varsigma = \frac{1}{T}\int_0^T |\mathbf{x}(t) - \mathbf{y}(t - d(t))|\, dt,$$

where the integration is over T the duration of the task. The global measure of distortion is very similar to the mean magnitude of $\mathbf{e}(t)$, with the difference being that the former has had the effect of the error due to the inevitable time delay between $\mathbf{x}(t)$ and $\mathbf{y}(t)$ removed. It can therefore be seen that there are many measures that can be used to describe the performance of the S-M system in tracking a 2-D target. The particular set of performance measures used to assess the performance of the S-M system depends on the type of information required on the system and whether a statistical analysis is to be performed.

The question whether 2-D tracking tasks can provide additional or more information on S-M function than 1-D tasks has not be explicitly investigated to date. Researchers have, however, compared tracking performance on 1-D and 2-D tasks in an effort to investigate the information processing capabilities of the S-M system and its ability to perform multiple tasks. If 2-D tasks are able to either provide information on S-M function which is additional to that provided by 1-D tasks, or are able to provide more sensitive measures of S-M function, it would be expected that this information would have its origin in any additional processing required to complete a 2-D task over that required for a 1-D task. Thus, if 2-D tasks are able to provide additional or more sensitive measures of S-M function than 1-D tasks the performance measures should indicate that the 2-D tasks put more stress on the S-M system and require more processing than the 1-D tasks. Researchers investigating the information processing capabilities of the S-M system by comparing 1-D and 2-D tracking performance have therefore indirectly investigated the differences in 1-D and 2-D tracking. These studies should provide support for the author's contention that 2-D tracking tasks can provide information on S-M function additional to that provided by 1-D tasks. It needs to be emphasised, however, that these previous studies were more concerned with how the addition of a second tasks effects S-M performance than whether 2-D tracking tasks could provide information different to that provided by 1-D tracking tasks.

To evaluate the information processing capabilities of the S-M system Garvey and Taylor (1959), Chernikoff *et al.* (1960), Todosiev *et al.* (1967), and Ziegler (1968) attempted to determine whether the S-M system could perform one tracking task as well as two tracking tasks. This evaluation was carried out by getting subjects to complete a 1-D and then a 2-D tracking task, considered by all of these researchers to be in effect a combination of two 1-D tracking tasks, and then evaluating whether tracking performance degraded on going from the 1-D (one task) to the 2-D tracking task (two tasks). The question of whether a 2-D tracking task can be considered to be a combination of two 1-D tasks or a single tracking task in its own right is addressed later. To determine if performance had degraded on going from the 1-D to the 2-D task, the above researchers all compared the MAE from the 1-D task to the MAE obtained from the coordinate in the 2-D task that was common to the 1-D task.

The tracking tasks were all of compensatory type with various forms of control dynamics installed after the subject's output transducer. These position (zero order), velocity (first order) and acceleration (second order) control systems inserted between the output transducer and the response signal output (see Fig. 8.2) alter the manner in which the subject must move the transducer to control the movement of the response pointer. For example, with the position control system (used for most tracking tasks) an output transducer displacement produces a proportional displacement in the response pointer, whereas with the velocity control system a movement of the output transducer in a particular direction increases the velocity of the response pointer in that direction and moving the output transducer in the opposite direction proportionally increases the response pointer velocity in the opposite direction. The control dynamics used by Ziegler (1968) were either position, velocity, acceleration, or third order control dynamics, while Garvey and Taylor (1959) used either position or acceleration control dynamics, Todosiev *et al.* (1967) used acceleration control dynamics, and Chernikoff *et al.* (1960) used position, velocity, and acceleration control dynamics. All of the above researchers used different techniques to produce the 1-D and 2-D target signals. In the study by Ziegler (1968) the same signals, consisting of the summation of two sinusoids, were used in each coordinate; Todosiev *et al.* (1967) state different random noise signals of the same bandwidth were used; Garvey and Taylor (1959) produced a 2-D signal by rotating the 1-D signal, formed by summing three sinusoids, 45 deg from the horizontal axis, and Chernikoff

et al. (1960) used two completely different sine waves. The results of Chernikoff *et al.* (1960) were probably weakened by the fact that a sine wave, which is highly predictable, was used as the tracking signal.

The results obtained from these studies were conflicting with Garvey and Taylor (1959) and Ziegler (1968) demonstrating a drop in tracking performance on going from a 1-D task to a 2-D task, while Chernikoff *et al.* (1960) and Todosiev *et al.* (1967) could find no drop in performance. Ziegler (1968) argues that if no difference is observed between 1-D and 2-D tracking then the 2-D task did not stress the S-M system sufficiently. On comparing 2-D and 1-D tracking performance, Ziegler (1968) determined that the 10% increase in tracking error, observed for the 2-D task with position control dynamics, was not significant while the increases observed for all the other forms of control were significant (the figures given within this paper seem to indicate that the increase in error for the position control dynamics was indeed significant). Ziegler (1968) therefore concluded that the task with position control dynamics was not sufficiently difficult to stress the S-M system whilst those tasks with higher order dynamics, which are more difficult, did sufficiently stress the system. The negative results of Chernikoff *et al.* (1960) and Todosiev *et al.* (1967) might attributed to the low bandwidth of their target signals in comparison to those of the other researchers and the fact that Chernikoff *et al.* (1960) used a predictable sine wave target signal. On evaluating the results from these studies and the type of tracking tasks used, there appears to be evidence to support the notion that tracking performance in the coordinate of a 2-D task that is common to a 1-D task is worse than that observed with the 1-D task.

In addition to the above researchers, others (Damos and Wickens, 1977; Gopher and Navon, 1980; Navon *et al.*, 1984; Fracker and Wickens, 1989) have more recently used 1-D tasks in conjunction with dual-axis tracking tasks and 2-D tracking tasks to determine how the additional information processing required by the latter two tasks is carried out and to determine how the S-M system performs when its information processing capabilities are stressed. Of particular relevance is the study by Navon *et al.* (1984). This study compared 1-D to 2-D pursuit tracking by evaluating whether performance on either task degraded when a second unrelated task (binary digit classification) was performed simultaneously. It was found that the second task competed for some of the resources required by the tracking tasks and there was a corresponding drop in performance for both tasks, however the observed drop in performance for the 2-D task was not significantly different to that observed for the 1-D task. Thus, there was no evidence that 2-D tracking was significantly more demanding than 1-D tracking. The experimental design of Navon's study may have contributed to this finding as it relied on the assumption that the extra resources that may be required to do the 2-D task come from that same pool of resources that the binary digit classification task and 1-D tracking task competed for. Also, if the difference between performance on 1-D and 2-D tasks is small and the subject allocates the majority of the resources to the tracking task when tracking the 1-D and 2-D targets, the binary digit classification task may not be sensitive enough to detect any difference.

Navon *et al.* (1984) and Fracker and Wickens (1989) also address the question of whether the S-M system treats a 2-D tracking task as a single task or two 1-D tasks. Fracker and Wickens (1989) concluded that a dual-axis task is a single task and that the tasks combined to form the dual-axis task can also be considered to be single, less difficult, tasks when performed separately. The complexity of the dual-axis task is directly related to the ease with which the subject can solve the problem of controlling the two component tasks at the same time. For a 2-D tracking task, with the same control dynamics in each coordinate, this problem is easy to solve with the subject attempting to minimise the radial error. To further support the argument that 2-D tracking tasks are single tasks, Fracker and Wickens (1989) also found that the subjects

controlled both axes almost always simultaneously, while for dual tasks with different dynamics the subjects frequently controlled the axes separately. Navon *et al.* (1984), however, could find no evidence to support the notion that 2-D tracking was a single task. In the study by Navon *et al.* (1984), random pursuit tracking and size tracking (controlling the size of a rectangle to match that of a target rectangle) tasks were used as heterogeneous or homogeneous task pairs in a dual task. The heterogeneous situation is where size tracking is performed in one coordinate while pursuit tracking is performed in the other and the homogeneous situation is where either size tracking or pursuit tracking is performed in both coordinates (i.e. in the homogeneous situation the tracking task is a 2-D tracking task). With the pursuit tracking task, the subject's response and target were displayed using rectangles and the object of the task was to minimise the distance between the bottom left vertices of the two rectangles. If 2-D tracking was a single task it was thought that the single axis performance on the tasks with homogeneous pairs would degrade less than with heterogeneous pairs. However, no difference was observed. Using the reasoning of Fracker and Wickens (1989), the findings of Navon *et al.* (1984) may be due to the fact that solving the problem of doing the two heterogeneous tasks simultaneously did not tax the human S-M system any more than doing two homogeneous tasks simultaneously. Overall, it is probably reasonable to accept the intuitive view that a 2-D tracking task is a single task and not two tasks. That is, when a subject tracks a 2-D target, they attempt to reduce the radial error rather than independently attempting to reduce the error in each coordinate.

Accepting the view that a 2-D tracking task is a single task, the question of whether tracking in 2-D is more demanding than tracking in 1-D is still unanswered, although there appears to be some evidence in support of this belief. It is also questionable whether comparing the tracking performance in the coordinate common to the 1-D and 2-D tasks is a valid method of determining whether 2-D tracking tasks are more demanding than 1-D tracking tasks. This is because if a 2-D task is constructed from two 1-D tasks of the same *difficulty*, the overall difficulty of the 2-D task is greater than that of the 1-D tasks (i.e. the average velocity and displacement of the 2-D task is greater than that of the 1-D task). As a result, it is not possible to distinguish whether any difference in performance on going from 1-D to 2-D is due to the 2-D nature of the task or the increased difficulty. Original work presented within Chapter 10 attempts to obtain a definitive answer to the question of whether 2-D tasks are more demanding than 1-D tasks and whether they can therefore provide different information on S-M function.

8.4 CHRISTCHURCH HOSPITAL'S SENSORY-MOTOR TEST BATTERY

The advantages of using quantitative tests to measure and analyse S-M function in clinical practice and research was recognised by staff at the Christchurch hospital in the late 1970s. A comprehensive computerised S-M test battery, based around a PDP11/34 computer and VT11 graphics display, was therefore developed and installed in the Neurology Department of Christchurch Hospital (Jones, 1986b; Jones *et al.*, 1989). This test battery was designed to quantify the integrity of S-M subsystems via tests of visuospatial function, strength, ballistic movements, and steadiness, and integrated S-M function via 1-D tracking tasks (Jones and Donaldson, 1981; Jones and Donaldson, 1986; Jones, 1986b). This system has had extensive use as both a clinical rehabilitation tool (Jones *et al.*, 1983; Croft and Jones, 1987) and research tool (Jones and Donaldson, 1981; Jones, 1986b; Jones *et al.*, 1986; Jones *et al.*, 1989; Jones and Donaldson, 1989; Jones *et al.*, 1990; Jones *et al.*, 1992).

Recently, it became clear that the PDP11-based system was physically reaching the end of its life, so it was decided to redevelop the battery on a more modern computer system able to

facilitate the future development of the system. This system upgrade also involved a complete redevelopment of the software. In this section a description of the hardware, software, features, and system design philosophy of the upgraded S-M test battery is given. The emphasis in the description is on those features of the upgraded battery which are of direct relevance to the 2-D tracking tasks incorporated into the upgraded battery by the author. A more general description of the upgraded S-M test battery can be found in Jones *et al.* (1993).

The upgraded S-M test battery consists of a number of tests which have been designed to measure component and integrated upper-limb S-M functions. The set of tests forming the battery can be broadly categorised into those used to assess visuospatial function, those used to assess the function of the motor subsystem, and tracking tasks. The tests used to assess the visuospatial components of the S-M system are the visual acuity, visual resolution, arrow comprehension (i.e. comprehending whether a flashing dot is on the head of an arrow or not), and static and dynamic perception test (i.e. determining the position of arrow with respect to a static or moving target). These tests assess the function of the visual system and its associated components by getting the subject to make a verbal response to a visual stimulus. The tests designed to measure the function of the motor components of the S-M system are the range of movement, grip strength, arm strength, ballistic movement, static steadiness, dynamic steadiness, and multifinger tapping (paced and unpaced with one finger, two fingers, and four fingers) tests. These tests require a minimal amount of visual processing and are basically designed to assess the the fundamental functions of the motor system as it relates to the upper-limbs. The final category of tests are the tracking tasks which assess integrated S-M function by placing substantial demands on both the visual and motor subsystems. The suite of tracking tasks consist of a series of 1-D pursuit tasks which require the subject to follow a target controlled by either a sinewave, a random signal, step functions, or a combination of the random signal and the step functions. With the 1-D tracking tasks there is the option of giving the subject a preview of the target signal by having the target signal descend from the top of the screen (see Fig. 1 Jones and Donaldson (1986)). The 2-D pursuit tracking tasks require the subject to follow a target controlled by either a random signal or a step function. No preview is given of the target signal. The 2-D tracking task developed by the author and the tapping task developed by Muir *et al.* (1992) are tests that were incorporated within the battery at the time it was upgraded. For detailed descriptions of the tests and performance measures used with the tests the interested reader is referred to Jones (1986b) for the original component tests, to Jones and Donaldson (1986) and Jones (1986b) for the 1-D tracking tasks, to Muir *et al.* (1992) for the tapping tests, and to §9.1 for the 2-D tracking tasks.

The target signals used for the 1-D and 2-D random pursuit tracking tasks are pregenerated. The random signals are produced by summing an appropriate number of harmonically related sinusoids of different phase. By this means it is possible to have several different pseudo-random signals that are non-periodic within the 120 s duration of the 2-D task, have a flat spectrum within a user specified bandwidth, have no components above this bandwidth, and whose spectrum (and that of the subject's response) can be accurately computed using the fast Fourier transform algorithm (Brigham, 1988). To satisfy the first criterion, the fundamental frequency of the signal must be less than 0.008 Hz (1/120 s). Since the target's position is updated at the vertical interrupt rate of the graphics display, 7241 samples (120 s x 60.34 Hz) are required in the two minute test interval. However, to satisfy the third criterion, the number of samples must be an integer power of two and, thus, a minimum of 8192 samples are required, giving a period of 143 s and a fundamental frequency of 0.007 Hz. The random signal generation program asks the user for the required signal bandwidth and then calculates the number of equal amplitude harmonics that must be summed together to give this bandwidth.

Each of these harmonics is assigned a phase equal to a uniformly distributed random number between 0 and 2π. A sine wave target can also be generated if a predictable target signal is required.

The new computerized S-M test battery has been developed to meet several design goals. A primary aim was to develop a system built around a current main-stream computer. It was therefore decided that the test battery would be best based around the low cost and readily available IBM family of personal computers. The system is currently built around a PC with an 80486 CPU running at 33 MHz, although the test battery has run quite adequately on a less powerful PC (e.g. 386 PC at 20 MHz). The system runs under DOS, requires 2 Mb of RAM and a co-processor is recommended. Another important aim was to physically separate the test and analysis functions. Thus the hardware comprises two high-resolution large screen (17 inch) colour monitors: one for the tracking display, and one for use by the assessor for task control and analysis. The two monitors are driven by WinSprint 100 graphic controllers (Artist Graphics Inc), in 640x480 VGA and 800x600 EGA modes respectively, and can be controlled simultaneously to produce dynamic colour graphics. The size of the display area on the tracking display monitor has been adjusted to achieve equal pixel dimensions in both the horizontal and vertical directions.

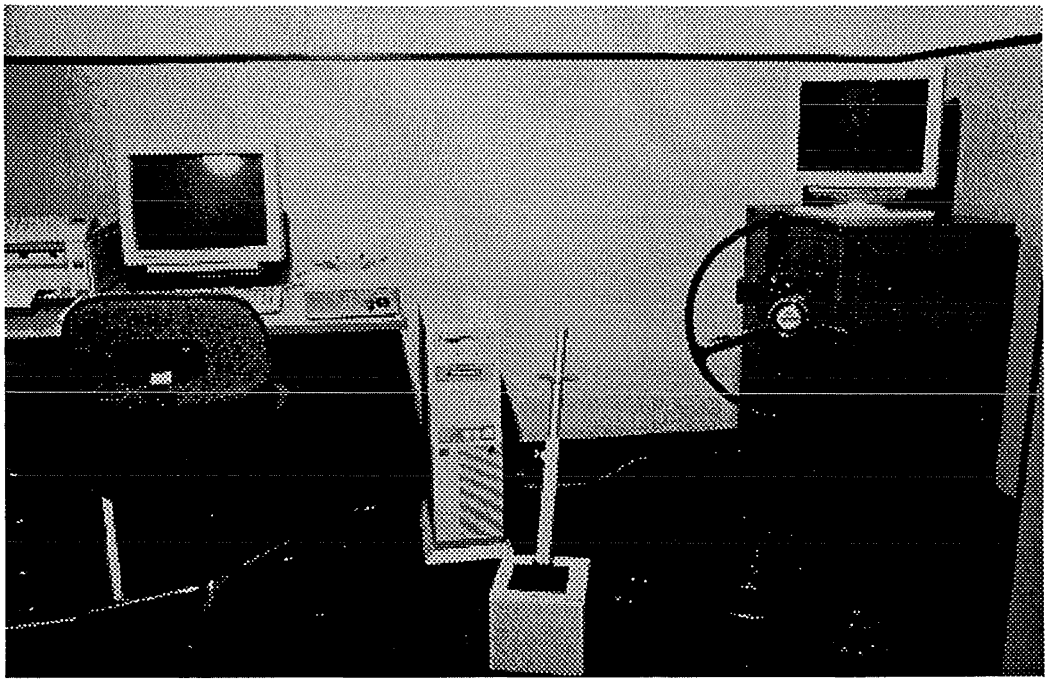

Figure 8.4 The S-M test battery comprising a 486/33 PC, dual colour monitors, floor-mounted joystick, steering wheel, and tapping sensor.

To measure the subject's S-M output, the upgraded S-M test battery uses a number of subject output transducers. These consist of a steering wheel (primarily used for the 1-D tracking tasks), a floor-mounted joystick, the computer's mouse, and a multifinger tapping sensor (see Fig. 8.4). The floor mounted joystick or the computer's mouse provide the subject's motor input for the 2-D tasks. The joystick has an adjustable height of 470 mm to 710 mm, measured from the bottom of its weighted base, and an angular range of movement of 10 deg in any direction from its vertical axis. Moving the joystick one degree produces

a corresponding 10 mm movement on the screen. Two orthogonal potentiometers detect the angular positions of the joystick and their output is sampled and digitised to 12 bits by the computer. Movement of the joystick can also be restricted to 1-D in either the lateral or fore-aft direction, corresponding to the horizontal and vertical directions on the screen respectively, if desired.

Considerable attention has been paid to designing a system which can be operated with minimum training. Thus the user interface is fully menu driven and incorporates a context-sensitive on-line help system. The system is also robust in terms of its tolerance to operator errors, with all commands and input data being checked for validity; default input data is available where appropriate. On completion of a test, the subject's test data is automatically stored on disk and analysed. Performance analyses are displayed in tabular and graphical form on the computer monitor and, if desired, as hard copy from a laser printer.

The tracking task software is structured around a central unit which produces the menus and calls the appropriate task control unit for either the 2-D, 1-D or component tasks. The task control units then call routines to produce the required task, analyse the results, and present a summary of the subject's performance. The routines that perform the data storage, data retrieval, and graphical displays of performance are common units, shared by all the tasks. The software for the S-M test battery has been written in Turbo Pascal with considerable use of object oriented code. Exceptions to this are the routines needed to drive the EGA stimulus display, which have been written in assembly language to achieve greater speed and, hence, jitter-free dynamic graphics.

The assembly language routines generate a display of the target and the subject's response marker by considering the video memory (configured in EGA mode) as four overlapping planes, each switchable (via a mask), and each capable of displaying the background colour and a single colour from a palette. To understand how these planes are formed and how the mask controls their display, it is necessary to note that each pixel on the monitor is allocated one nibble in the video memory (i.e. 4 bits, sufficient represent sixteen colours). To display a pixel on the monitor the contents of the nibble are ANDed with the mask, which is another nibble which in the normal configuration has the contents 1111, to give a binary representation of the colour to be displayed at the pixel. This binary representation is then looked up in a colour table (palette) which defines how the electron guns must be controlled to produce the required colour on the monitor. For example, if the nibble in video memory at the current location has the contents 1101, the mask has the contents 1000, and in the palette colour 1000 is defined to be white, the colour white is displayed on the monitor at the current location as $1101 \cdot 1000 = 1000 =$ white. By only ever setting one bit in the mask it can be seen that it is possible to separate the video memory into four planes which can be switched on and off via the mask.

To display, jitter free, the target and pointer for the tracking tasks, two planes are used for the purpose of displaying the target (e.g. those defined by setting the mask to either 1000 or 0010), with the remaining two planes being used to display the subject's pointer (i.e. 0100 and 0001). In the palette the colours 1000 and 0010 are then defined for example to be green and the colours 0100 and 0001 are defined to be red. By setting the bit plane mask to 1100 the current green target and the current red subject's pointer can be displayed on the monitor, while the next view of the target and subject's pointer, in their new position, are being drawn on the remaining two undisplayed planes (i.e. when the mask is set to 1100 it is possible to change the last two bits of each nibble in video memory without these changes being observed on the monitor). The bit plane mask is toggled between 1100 and 0011 each time the

computer receives a vertical synchronization interrupt from the graphics controller indicating the completion of a raster and, hence, indicating that the planes with the updated target and response positions can be displayed. Through a combination of a high update rate of 57.26 Hz (i.e. the vertical interrupt frequency) and dual display buffers, it has been possible to obtain an extremely smooth dynamic colour graphics display.

The upgraded S-M test battery is now getting extensive use in the clinical area where it is routinely used as a central part of the off-road section of a driving assessment and training program. The primary focus of this program is to determine and recommend whether people who have some form of neurological damage are capable of driving safely The versatility and validity of the battery in the S-M research arena has been demonstrated by the measurement of central executive deficits in Parkinson's disease (Dalrymple Alford *et al.*, 1994), multifinger tapping rhythm in Parkinson's disease (Muir *et al.*, 1992), 2-D step tracking deficits in Parkinson's disease (Jones and Watson, 1993)(see §10.1 also), and the differences between 1-D and 2-D tracking in both normal and parkinsonian subjects. The results of the author's comparison of 1-D and 2-D tracking performance and step tracking with normal and Parkinsonian subjects are presented in Chapter 10.

CHAPTER 9

TWO 2-D TRACKING TASKS AND TRACKING TASK TIME DELAY ESTIMATION

The original S-M test battery at Christchurch Hospital consisted of 1-D tracking tasks and a number of tests designed to measure component S-M function. None of these existing tests explicitly investigated the ability of a subject's S-M system to function in a multidimensional environment, similar to that in which it must function every day. It was felt that multidimensional tasks might be able to provide new or more sensitive measures of S-M function because of the spatial comprehension and additional processing that should be required to complete the task. For this reason the author developed two 2-D tracking tasks which have been incorporated within the Christchurch Hospital S-M test battery.

One of the primary measures of performance used with the 1-D tracking tasks and the 2-D tracking tasks is the estimate of the time delay between the signals describing the target motion and the subject's response. The 1-D tracking tasks within the original S-M test battery used an adapted spatial domain cross-correlation technique, described in §9.2.3, to provide an estimate of the time delay. The time delay estimates provided by this technique were found, however, to be inaccurate in many situations. The author has therefore conducted a study to determine why the adapted cross-correlation technique is so unreliable. A study evaluating the accuracy of some of the many techniques that can be used to provide the time delay has also been conducted to determine which technique can most reliably provide accurate average time delay estimates in the tracking situation.

This chapter contains the author's original research in the S-M area which relates to the development of the 2-D tracking tasks. In the first section of this chapter the two 2-D tracking tasks developed by the author are described along with the performance measures used with the tasks. This is followed in §9.2 by a review of some of the various techniques that can be used to estimate the time delay that exists between the target and subject's response signals. The limitations of the various techniques are described and the results obtained from the study, designed to assess the accuracy of the various time delay estimation techniques, are presented.

9.1 TWO 2-D TRACKING TASKS

Two 2-D tracking tasks have been developed and incorporated within the upgraded computerised S-M test battery at Christchurch Hospital. The two tasks developed were a 2-D random pursuit tracking task and a 2-D step pursuit tracking task. The 2-D random pursuit tracking task assesses the ability of the subject to track a smoothly changing random signal by using a closed-loop tracking strategy (i.e. the subject continuously assesses the current target and response pointer position and updates the tracking response to minimise the error between the

target and the response pointer). The 2-D step tracking task was designed to measure the other end of the S-M spectrum by getting the subject to accurately perform ballistic movements. When the S-M system performs ballistic movements it operates in an open-loop manner and uses a preprogrammed movement plan to implement the response (i.e. no feedback is used during the execution of the movement to correct for any error). This task has been included within the test battery as past research has shown that some neurological disorders, such as Parkinson's disease, impair the ability of the subject to perform open-loop movements (Flowers, 1976). Although the concept of a 2-D random pursuit tracking task is not original, as the references cited in §8.3.4 indicate, the 2-D step pursuit tracking task is original in both concept and design. The 2-D tasks were added to the upgraded battery for the reasons outlined above and because they are likely to stress the S-M system to a greater extent than 1-D tasks (both from visuospatial and motor planning points of view) and are thus likely to be more sensitive in terms of allowing subtle differences in S-M function to be detected. The two 2-D tracking tasks developed by the author are now described.

9.1.1 2-D Random Pursuit Tracking Task

The 2-D random pursuit tracking task requires a subject to follow the movement of a target square with a 10 mm $\times$ 10 mm cross (see Fig. 9.1(a)) controlled by a joystick or mouse. Both the horizontal and vertical positions of the target are defined by similar smoothly changing random signals, with the vertical position simply being a phase shifted version of the target's horizontal position. The resulting target movement has been confined to a 139 mm $\times$ 139 mm square area centrally situated on the monitor. The system described in §8.4 which is used to generate random target signals is used to pregenerate the random signals required for this task. Alternative sinusoidal target waveforms can also be pregenerated using this system if a predictable target signal is required.

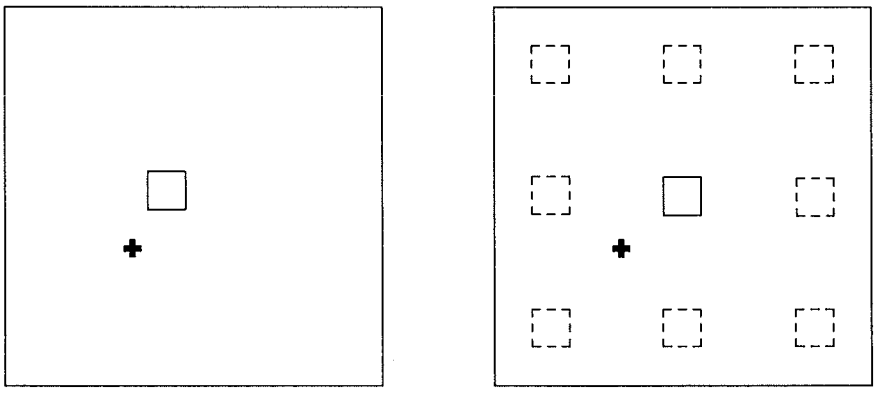

Figure 9.1 Schematics of visual displays for (a) the 2-D random pursuit tracking tasks, and (b) the 2-D step pursuit tracking task.

To begin the 2-D random pursuit tracking task, the assessor selects the task from the menu of possible 2-D tasks. A dialogue box then confirms that the standard 2-D random pursuit tracking task is required giving a target square of 10 mm $\times$ 10 mm, a task duration of 120 s, and a target signal bandwidth of 0.20 Hz. The default target signal has a mean displacement of 31 mm, maximum displacement of 86 mm, mean velocity of 24 mm/s, and maximum velocity of 68 mm/s. If a variation of the standard task is requested, a form appears on the screen,

on to which the assessor can enter a target box size not exceeding 44 mm $\times$ 44 mm, a test duration not exceeding 140 s, and can select any of the available target signals. In addition to these parameters, the assessor also has the ability to specify the position within the target signal at which the test begins and whether the target movement is to be restricted to either the vertical or horizontal direction. Once the task parameters have been specified, a set of standard instructions appears on the control monitor for the assessor to read to the subject. On the task display monitor a blue target box is displayed along with the response cross. The subject is then able to obtain a feel for the joystick/cross dynamics by moving the joystick and observing the corresponding cross movement. To commence tracking the subject must place the cross within the blue target square for 2 s. The target box then turns green, indicating to the subject that the test has started and, after a 4 s delay, the target begins to move. The subject's response is sampled at 28.63 Hz (i.e. half the vertical interrupt frequency).

Measures of the subject's performance are automatically calculated once the task has been completed and the raw data stored. The primary error measure used to evaluate tracking performance is the mean absolute error (MAE), the mean value of the absolute distance/error (AE) between the centre of the target box and the centre of the response cross at each sample. For the 2-D tasks, the AE and its mean values are computed in both the horizontal and vertical coordinates ($MAE2D_X$ and $MAE2D_Y$ respectively) and the resultant ($MAE2D_R = mean\left\{\sqrt{AE_X^2 + AE_Y^2}\right\}$). The mean AE over six second epochs is also calculated, so that a histogram of performance during the task can be constructed. Error histograms are useful for detecting and quantifying within-task variations in performance due to learning, fatigue, and lapses in concentration. Care must, however, be taken in reaching such conclusions as the task *difficulty* for random targets (cf. sinusoidal targets) will generally vary between epochs (Jones and Donaldson, 1986). Performance biases in terms of errors to the top *vs* bottom and left *vs* right of the target, and time-on-target, expressed as a percentage of test duration, are also computed. The time-on-target performance measure is considered a poor measure for the reasons outlined in §8.3.2 but was included with this task as colleagues of the author had a need for such a measure (Dalrymple Alford *et al.*, 1994).

To provide information on the tracking strategy adopted by the subject, the lead or lag ($Lag2D_X$, $Lag2D_Y$) and gain ($Gain2D_X$, $Gain2D_Y$) of the subject's response are computed for each coordinate. As described in §8.3.3 the gain provides an indication of whether the subject tends to either over- or under-shoot the target and the time delay estimation provides an indication as to the size of the information processing delays in the S-M system. The time delay estimate is computed using a Fourier domain cross-correlation technique as this technique was found to be one of the most accurate techniques. The results from the study used to determine which time delay estimation technique could most accurately and reliably compute the time delays in the tracking situation are presented in §9.2. All of the above error measures are displayed in tabular form to the examiner, with a histogram of the MAE in each epoch also being available. In addition, graphs of the phase difference between the target and response signals and of the spectral magnitude of the target and response signals can be displayed.

9.1.2 2-D Step Tracking

To allow examination of ballistic movements, a 2-D step pursuit tracking task has been developed. In this task the subject must quickly and accurately move the cross from within a starting square to within a target that appears on the screen with temporal and spatial unpredictability. As shown in Fig. 9.1 (b), the display for the step tracking task comprises eight 10 mm $\times$ 10 mm

squares arranged around a centrally situated 10 mm × 10 mm square. The centres of the eight surrounding target squares are positioned at the vertices and midway along the perimeter of an imaginary 100 mm × 100 mm square centred on the central square.

To initiate the task, the subject places the cross within the perimeter of the red central target. The target turns green to indicate to the subject that the task has started. After either a 2, 3, 4 or 5 s delay, one of the surrounding blue targets turns green (and the central target turns blue) and the subject moves the cross to within the green target square as quickly and as accurately as possible. After a further delay of either 4, 5, 6 or 7 s the central target turns green and the subject moves the cross back. Note that in the latter movement the position of the target is known and so there is only temporal uncertainty. The task continues with the green target square alternating between the central square and any one of the surrounding squares until ten outward and ten return targets have been presented to the subject. During the course of the task the subject's response is sampled at a frequency of 60.34 Hz, well above the Nyquist rate of approximately 18 Hz (Jones and Donaldson, 1986).

To enable valid comparison of results between subjects, the step target sequence is the same for every run, with the exception of a randomly selected phase shift. The phase shift, obtained by altering the starting position within a fixed array containing the sequence of targets and delays, maintains the spatial and temporal unpredictability of the task if the subject needs to carry out the task more than once.

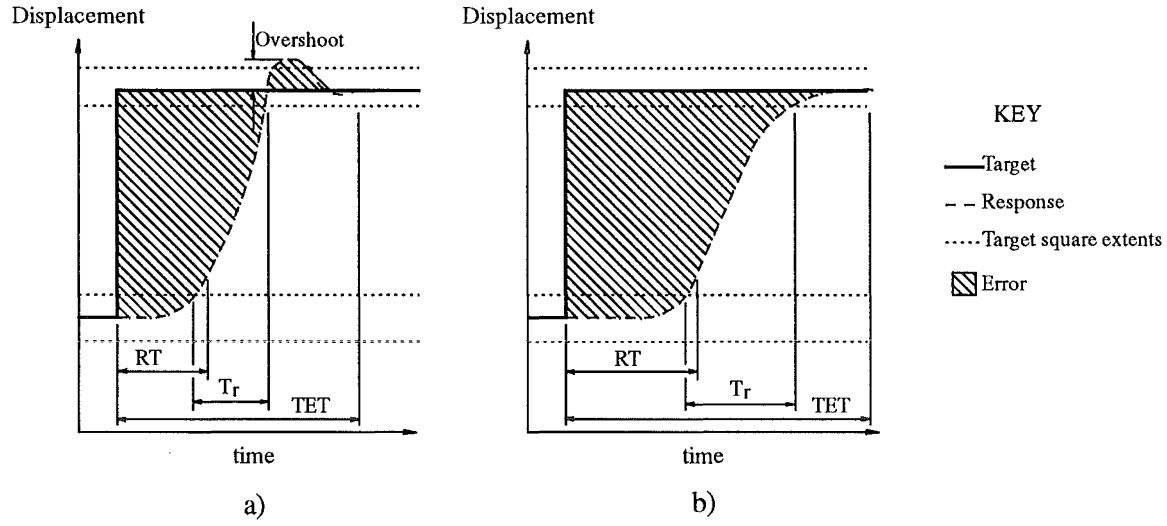

Figure 9.2 Typical step tracking responses and performance parameters used for (a) an underdamped response, and (b) an overdamped response. RT = reaction time, T_r = rise time, TET = target entry time.

The responses for each individual target square are separated into their horizontal and vertical components and several measures are computed independently for each. The first measure is the *reaction time*, defined as the time delay between the target square turning green and when the subject moves the entire cross outside the previous target square (Fig. 9.2). If the reaction time is less than 200 ms, the subject is considered to have anticipated the target movement and the response measures are invalidated. The *rise time* is the time taken for the subject to move from 10 % to 90 % of the magnitude of the primary movement (i.e. the movement beginning with the initiation of movement and ending when the cross's velocity first returns to zero). The *overshoot* is the distance between the centre of the cross at the end of the primary movement and the centre of the green target square, expressed as a percentage of the distance to the target square (i.e. 50 mm) and with a negative value indicating undershoot of

the target. The *target entry time* is the interval between the target square turning green and the entire cross entering, and remaining settled, within the target square. An overall *error* measure is computed by determining the area between the ideal response (i.e. a step) and the subject's response (see Fig. 9.2); this is normalized with respect to the area calculated if the subject did not respond at all to the target.

To reduce the results into a more manageable form, the mean value of the five minimum reaction times, rise times, overshoots, target entry times, and errors were computed independently for the inward and outward movements. Following the first clinical study using this task (see §10.1), it was found that each of the reduced performance measures was usually computed from a set of five different targets and so it was not valid to compare the different reduced measures to deduce information about the tracking strategy used. To overcome this problem the reduced measures, for each performance measure, are now computed by taking the median values of the results obtained from: eight movements to the eight different surrounding target squares, and eight movements from the eight different surrounding squares to the central target square. Only eight out of the ten inward or outward movements are used in the computation of the median performance measures as two of the ten movements represent movements to a repeated target. Left-right and top-bottom (i.e. fore-aft on joystick) performance biases are also calculated. To provide information on the tracking strategy adopted by the subject, graphical displays of the subject's trajectory to each target, in both horizontal and vertical directions, are also available.

9.2 COMPUTATION OF A SUBJECT'S AVERAGE TIME DELAY IN PURSUIT TRACKING

It was was indicated in §8.3.3 that the time delay between the target signal and a subject's response signal can provide valuable information on the information processing delays in the S-M system. To compute the time delay, the cross-correlation of the target and response signals is usually calculated in either the spatial or frequency domain. The location of the maximum of the cross-correlation function then provides an estimate of the average time delay between the two signals. It is shown in this section that the finite length of the two signals biases the time delay estimate towards a delay of zero. To overcome this bias an adapted version of the spatial domain cross-correlation technique was used to compute the tracking delays within the original S-M test battery (Jones, 1986b). The present author found the technique to be inaccurate. The inaccuracy with the adapted cross-correlation technique prompted the author to investigate why this technique provides such inaccurate time delay estimates. The author was also interested in determining which of several time delay estimation techniques could provide the most accurate estimate of the time delay in the tracking situation. To determine which method could provide the most accurate time delay estimates in the tracking situation a study of delay estimation accuracy has been conducted using simulated responses.

Five time delay estimation techniques were used to determine which technique could provide the most accurate tracking delay estimations. These consisted of three spatial domain and two Fourier domain techniques: spatial domain cross-correlation (SCC), adapted spatial domain cross-correlation (ACC), least squares time delay estimation (LSE), frequency domain cross-correlation (FCC), and phase shift delay estimation (PSE). In §9.2.1 the literature describing previous studies, which have compared various time delay estimation techniques by computer simulation, is reviewed. In the sections following this review, each of the above time delay estimation techniques is then described in detail along with each technique's strengths

and weaknesses. Of particular note are the discussions in §9.2.2.2 and §9.2.3.2 describing the bias in the standard cross-correlation technique and the reason for the inaccuracy with the adapted cross-correlation technique respectively. As the target and response signals are discrete signals the descriptions of the techniques are generally in terms of discrete signals, although limitations of the techniques are explained in terms of continuous signals for clarity. In §9.2.6 the methods used by the author to compare the accuracy of the time delay estimation techniques, in the tracking situation, is described and the results from the experimental study are presented and discussed.

9.2.1 Review

Before describing the time delay estimation techniques compared in the author's study, it is beneficial to briefly review the methods and findings of other researchers who have conducted similar studies. In recent years a number of researchers have used experimental studies to compare the accuracy and robustness of a wide range of time delay estimation schemes. The interest in such techniques stems from the need in many areas of research, technology, and industry to be able to measure the time delays that exist between contaminated versions of similar signals. Some of the common areas of application for time delay estimation techniques are: radar, sonar, process control, velocity and distance measurement, pitch estimation in speech, propagation delays in nerves, the characteristics of multipath environments, and stereo-vision (Fertner and Sjölund, 1986; Avitzour, 1991; Jacovitti and Scarano, 1993). To determine the time delays that occur in these areas, techniques such as standard cross-correlation (Scarbrough *et al.*, 1981; Bedini *et al.*, 1986; Fertner and Sjölund, 1986; Avitzour, 1991; Jacovitti and Scarano, 1993), generalised cross-correlation (the signals are pre-filtered or emphasised prior to cross-correlating) (Scarbrough *et al.*, 1981; Bedini *et al.*, 1986; Avitzour, 1991), hybrid sign and polarity coincidence correlation (adapted versions of standard cross-correlation designed to reduce computation) (Fertner and Sjölund, 1986), least squares type estimators (Fertner and Sjölund, 1986; Jacovitti and Scarano, 1993), the Meyr-Spies Method (Fertner and Sjölund, 1986), and maximum likelihood estimators (Avitzour, 1991; Jacovitti and Scarano, 1993) are used. To compare the accuracy of the above techniques researchers usually produce a signal that has the same characteristics (e.g. bandwidth) as the signals observed in the application of interest. A version of this signal is then delayed and Gaussian distributed noise is then added to the resulting signal to give a delayed contaminated version of the original signal. In many cases Gaussian distributed noise is also added to the original signal (Scarbrough *et al.*, 1981; Bedini *et al.*, 1986; Fertner and Sjölund, 1986; Avitzour, 1991; Jacovitti and Scarano, 1993). These signals are then given to the various time delay estimation techniques and the ability of the various techniques to estimate the known time delay is then assessed for various SNRs and when the number of samples used to compute the delay is varied (Scarbrough *et al.*, 1981; Bedini *et al.*, 1986; Fertner and Sjölund, 1986; Avitzour, 1991; Jacovitti and Scarano, 1993).

The consensus from these studies appears to be that the standard cross-correlation technique provides the best estimate of the time delay at most SNRs (Scarbrough *et al.*, 1981; Fertner and Sjölund, 1986; Avitzour, 1991). The standard cross-correlation technique has also been found to maintain this time delay estimation superiority as the number of signal samples used to compute the time delay is varied. However, when the SNR becomes quite high (i.e. ≥ 30 dB), time delay estimates provided by the least squares techniques have been found to be superior to those provided by standard cross-correlation techniques. This is because, in the absence of contamination, the least squares technique can provide an exact estimate of the time delay

whereas the standard technique cannot (see §9.2.2.2) (Fertner and Sjölund, 1986; Jacovitti and Scarano, 1993). It has also been shown that when the spectral characteristics of the original signals are known *a priori*, in some instances the generalised cross-correlation methods can be used to provide time delay estimates superior to those provided by the standard cross-correlation technique (Scarbrough *et al.*, 1981; Bedini *et al.*, 1986). None of the above studies have directly compared the time delay estimation accuracy of the five techniques compared in the author's study and to date no one appears to have evaluated the accuracy of time delay estimation techniques in the context of tracking tasks. The ACC technique also appears to be a novel technique for estimating the time delay, exclusive to the Christchurch Hospital research group, as to date no reference to a technique of this type has been found in the literature. Thus, the author's study of time delay estimation accuracy in the context of tracking and the limitations of the ACC technique are significant original contributions to the fields of S-M research and time delay estimation.

9.2.2 Standard Spatial Domain Cross-Correlation

9.2.2.1 Description of Technique

To calculate the time delay between two signals using the SCC technique the cross-correlation of the two signals is first computed. The discrete cross-correlation $z_{rh}[i]$ of two signals $r[n]$ and $h[n]$ which are nonzero over the time interval $N\Delta t$ can be calculated in the time domain by using the expression (Press *et al.*, 1986; Brigham, 1988)

$$z_{rh}[i] = \frac{1}{N}\sum_{n=0}^{N-1} r[n]h[n+i], \qquad i = -(N-1) \text{ to } N-1. \tag{9.1}$$

The above expression should be compared to Eqn.(2.53) which describes the cross-correlation of two continuous signals. The two signals are only nonzero over the time interval $N\Delta t$ because in practice it is only possible to record real life signals for a finite length of time and digital computers, which are usually used to compute the cross-correlation, can only store and process a finite amount of data. In most instances $r[n]$ and $h[n]$ are sampled sections of much larger and often continuous signals. For large N Eqn.(9.1) can be rewritten using expected value notation (Oppenheim and Schafer, 1975; Bendat and Pierslo, 1980) in the form

$$z_{rh}[i] = E\{r[n]h[n+i]\}, \qquad i = -(N-1) \text{ to } N-1,$$

where E denotes the expected value. In the situation where the signals to be cross-correlated consist of a signal and a noisy delayed version of the same signal, it is possible to express the delayed signal in the form $h[n] = r[n+n_\tau] + c[n]$ where n_τ is the delay. Assuming that the contamination is zero mean Gaussian distributed, as is commonly done in cross-correlation analysis (Scarbrough *et al.*, 1981; Bedini *et al.*, 1986; Fertner and Sjölund, 1986), it is possible to express the cross-correlation in the form

$$\begin{aligned} z_{rh}[i] &= E\{r[n](r[n+n_\tau+i] + c[n+i])\}, \; i = -(N-1) \, to \, N-1 \\ &= E\{r[n]r[n+n_\tau+i]\} + E\{r[n]c[n+i]\} \\ &= E\{r[n]r[n+n_\tau+i]\}. \end{aligned} \tag{9.2}$$

The expectation of $r[n]c[n+i]$ is zero for all values of i as a consequence of the noise being zero mean and Gaussian distributed. Provided $r[n]$ is not periodic with a period less than

the time interval $N\Delta t$, $z_{\tau h}[i]$ has a maximum when $i = -n_\tau$ since the two signals being cross-correlated are exactly aligned. As a result, the time delay present in a system can be determined by simply establishing for which value of time the cross-correlation of the system's input and output has a maximum. By fitting a suitable continuous function (e.g. a parabola) to the cross-correlation samples about the maximum it is possible to obtain sub-sample accuracy (Jones, 1986b; Moddemeijer, 1991; Jacovitti and Scarano, 1993). However, in the study described in §9.2.6 the delays were calculated only to the accuracy of the nearest sample.

9.2.2.2 Non-Ideal Characteristic of the Technique

In some situations where two signals have been cross-correlated in order to determine the time delay between them, the form of the signals can produce estimates of the error which are biased towards $n_\tau = 0$ (i.e. the SCC technique can underestimate the delay). The bias in the estimated delay, which is observed in practice, can be primarily attributed to the finite length of the signals used to compute the SCC. This bias in the computed estimate of the time delay was the motivation for developing the adapted cross-correlation technique that was used in the original S-M test battery and which is described in §9.2.3. The manner in which the finite length of the signals biases the results is now explained using the conceptual viewpoint of those researchers who originally devised the adapted cross-correlation technique. The author's attempt to obtain a more rigorous mathematical description of the phenomenon can be found in Appendix A. It should be noted that to date the author has been unable to find a detailed description of this phenomenon in the literature and therefore the description here is the author's own.

During the process of computing the SCC the number of non-zero terms in the summation in Eqn.(9.1) in effect decreases by one for each incremental increase in $|i|$, i.e. with each shift another sample of the shifted signal is no longer aligned with the nonzero samples of the other signal. In the case where the signals being cross-correlated are sections of a much larger signal (i.e. the larger signal is nonzero in a time interval $t >> N\Delta t$) the incremental decrease in the number of nonzero samples in the summation produces a cross-correlation function that has a form which is very similar to the product of a triangular window $w_{tri}[i] = ((N-1) - |i|)/(N-1)$ and the cross-correlation function that would have been computed if the larger signal was used to compute the cross-correlation (Beaucamp, 1973). It is shown in Appendix A that, for monochromatic signals that are truncated by a window having a length greater than the period of the signal and for multispectral signals with significant magnitude high frequency components, the cross-correlation function obtained in practice can be adequately described by the product of the triangular window function and the signal's analytic cross-correlation function. When a sinusoidal signal is multiplied by a triangular function it can be seen, as is shown in Fig. 9.3, that the maximum of the product is not aligned with that of the original signal. The effect of the multiplication is to shift the maximum in the direction of the apex of the triangular function. Thus, in the situation where the cross-correlation function is being computed in order to find the time delay in a system, the finite length of the signals used to compute the cross-correlation function result in a estimate of the delay which is biased towards $t = 0$. Experience has shown that the size of this bias is larger for low frequency sinusoids than for high frequency sinusoids. This is to be expected because as the number of periods used to compute the cross-correlation increases, the computed cross-correlation should become a more accurate approximation of the analytic cross-correlation.

The shifts due to above phenomenom can be quite severe if the signals being correlated have a significant dc component. The cross-correlation of two dc signals computed by Eqn.(9.1)

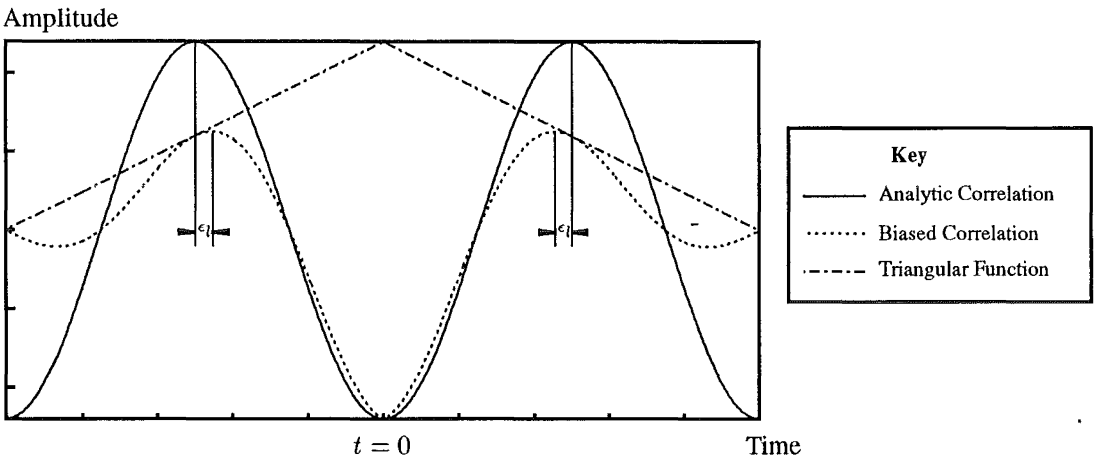

Figure 9.3 The bias, ϵ_l, that occurs in the position of the maxima of the cross-correlation function when the cross-correlation is computed using truncated time domain signals. Shown in this figure are: the analytic cross-correlation function of two infinitely long sinusoids, the triangular function that weights the analytic cross-correlation in the truncated case, and the resulting cross-correlation function with the maxima shifted towards $t = 0$

is simply a triangular function of width $(2N - 1)\Delta t$ which has a maximum at $t = 0$. In the worst case, where the dc components of the signals being correlated are much larger than the other spectral components, the cross-correlation is dominated by the triangular function arising from the dc components. The maximum of the cross-correlation in this situation would more than likely be at $t = 0$, even if there was a significant time delay between the two signals being correlated. When the dc components are not large enough to completely obscure the cross-correlation they can still move the maximum of the cross-correlation towards $t = 0$, thereby biasing the predicted time delay in this direction. The above effects indicate that it is advantageous to remove the dc components from any signals before their cross-correlation is computed via the SCC technique. In the study described in §9.2.6 the signals had their significant dc components removed before their cross-correlation was calculated. .

The inherent shift in the SCC technique prompted the development of the adapted cross-correlation technique. The adapted cross-correlation technique was designed so that the number of samples in the summation did not decrease with each shift. It was hoped that the adapted technique, which is discussed in §9.2.3, would remove any bias in the calculated time delays towards $t = 0$.

9.2.2.3 Calculation of Time Delay via Frequency Domain

The time delay between two signals can be computed via the frequency domain by using the frequency domain cross-correlation technique (FCC). This technique is equivalent to the SCC technique except that the calculation of the cross-correlation is carried out via the frequency domain using Eqn.(2.54). The resulting cross-spectrum is then inverse Fourier transformed to give the cross-correlation function and the time delay is found in the usual manner by searching for the maximum of the cross-correlation function. Before computing the signal's spectra it is necessary to zero pad the signals out to 2N since the cross-correlation ($z_{rh}[n]$), which is the inverse Fourier transform of the cross spectrum, is of length 2N-1. This technique of computing the cross-correlation induces the same bias in the computed time delay as the SCC technique. As with the SCC technique it is advisable to remove the dc component from the signals before

computing their spectra. Computing the cross-correlation function via the frequency domain is much simpler and computationally more efficient (if FFT algorithms are used) than computing the cross-correlation function directly in the spatial domain.

9.2.3 Adapted Spatial Domain Cross-Correlation

9.2.3.1 Description of Technique

To overcome the bias towards $n_\tau = 0$ that occurs with the standard spatial domain cross-correlation when finite length signals are correlated, an adapted version of the SCC technique was devised and incorporated within the original S-M test battery. With this adapted technique, known here as adapted spatial domain cross-correlation (ACC), the number of non-zero samples being used to compute the cross-correlation is kept constant for each increment in the index $|i|$. To demonstrate how this is accomplished, assume that the two signals to be correlated, $r[n]$ and $h[n]$, are known over the time intervals $M\Delta t$ and $N\Delta t$ respectively for $M > N$. If the magnitude of the shift between the two signals is at most $n_\tau \Delta t$, then M must be greater than or equal to $N + 2n_\tau$. In fact if the maximum of the ACC function is to be determined with certainty it is necessary to set $M = N + 2N_\Gamma$, where $N_\Gamma \gg n_\tau$. The shift between two signals which satisfy the above requirements can then be found by calculating the ACC, using the expression

$$\tilde{z}_{rh}[i] = \frac{1}{N} \sum_{n=0}^{N-1} r[n]h[n+i], \quad i = -N_\Gamma \text{ to } N_\Gamma. \tag{9.3}$$

As with the SCC technique described in §9.2.2, the shift between the two signals can be found by determining for which value of i the maximum of $\tilde{z}_{rh}[i]$ occurs.

It can be seen from Eqn.(9.3) that whenever $h[n]$ is shifted relative to $r[n]$, the number of non-zero terms in each summation does not change since the actual samples pertaining to $r[n]$ are always available. Thus, the the number of non-zero products in the correlation computation remains constant with each shift and the bias associated with the SCC technique should be eliminated. The ACC technique also has the beneficial characteristic that it is not necessary to remove the dc component from the signals before computing the correlation. This is because the dc components of the signals should add only a dc offset to the ACC function and not affect the position of the maximum of the cross-correlation. In the author's study N was set to 1024 and M was set to 2048.

9.2.3.2 Non-Ideal Characteristic of the Technique

With the ACC technique it was found that, as the simulation results show in §9.2.6, even in the case where an uncontaminated delayed and an undelayed version of the same signal were correlated the error in the calculated delay could be quite severe. The magnitude of the error in the estimated delay would also vary markedly depending on the form of the signals used to compute the correlation. The fragile nature of the accuracy of the ACC technique, which was originally devised to overcome the small error in the computed time delay that occurs with the SCC technique, was of concern to the author. For this reason the following analysis was performed to determine the origin of the large time delay estimation error that can occur with the ACC technique.

In order to aid the reader in understanding the origins of the error which arises with the ACC technique, an overview of the source of this error is first provided and then the mathematical

expression which describes the error is derived. The ACC technique can be thought of most simply as a SCC technique which operates on two signals that have been truncated to two different lengths by two rectangular windows of differing length. To enable the source of the ACC error to be described, consider the most straightforward case where two in-phase cosine waves, $s_1(t)$ and $s_2(t)$, are being correlated where $s_1(t)$ has been truncated by a window $w_1(t)$ of length l_1 and $s_2(t)$ has been truncated by a window $w_2(t)$ of length l_2. If the form of the spectra of the signals $s_1(t)w_1(t)$ and $s_2(t)w_2(t)$ are now considered along with their cross-spectrum, the origins of the ACC delay estimation error can be determined.

The spectra of the truncated signals $s_1(t)w_1(t)$ and $s_2(t)w_2(t)$ are the convolution of the spectrum of the cosine wave with that of the respective window function. Consequently, if the cosine wave has a frequency u_s the spectrum of $s_1(t)w_1(t)$ is simply the superposition of $W_1(u-u_s)$ and $W_1(u+u_s)$ and likewise the spectrum of $s_2(t)w_2(t)$ is the superposition of $W_2(u-u_s)$ and $W_2(u+u_s)$ (see Fig. 9.4 (c) and (f) respectively). The spectral magnitude of these windowing functions is described by sinc functions (i.e. $\sin(2\pi u)/2\pi u$) and, as the origin can be arbitrarily chosen to lie at the beginning of both windows, the spectral phase of the window functions is given by linear functions of the form $\pi l u$, where l is the length of the appropriate window. If the form of $S_1(u) \odot W_1(u)$ at $u = u_s$ is now considered it can be seen that $S_1(u) \odot W_1(u)$ is likely to have non-zero phase at $u = u_s$. This is because $W_1(u+u_s)$ is likely to have non-zero phase at $u = u_s$ and the summation of $W_1(u+u_s)$ and $W_1(u-u_s)$ at $u = u_s$ is therefore likely to produce a phasor with non-zero phase (see Fig. 9.4 (c)). The size of this non-zero phase term at $u = u_s$ is dependent on the window length. For the purpose of this argument let the phase at $u = u_s$ be θ_{l_1}, where the subscript denotes the dependence of this phase term on the length of the window. Similarly, at $u = u_s$ $S_2(u) \odot W_2(u)$ has a phase term θ_{l_2}. Now if the cross-spectrum, given by the expression

$$Z_{S_1S_2}(u) = (S_1(u) \odot W_1(u))(S_2(u) \odot W_2(u))^*, \tag{9.4}$$

is computed it can be seen that at $u = u_s$ the phase, θ_{l_D}, is equal to $\theta_{l_1} - \theta_{l_2}$ and at $u = -u_s$ the phase, $-\theta_{l_D}$, is equal to $\theta_{l_2} - \theta_{l_1}$ (see Fig. 9.4 (g)). These non-zero phase terms in the cross-spectrum of the two signals are the origin of the error that arises when the ACC technique is used to compute the time delay. This error arises because the ACC technique in effect uses two different length windows to truncate the signals and this results in θ_{l_1} and θ_{l_2} having different values which in turn produces non-zero phase terms in the cross-spectrum at $u = u_s$ and $u = -u_s$ (i.e. $\theta_{l_1} - \theta_{l_2} \neq 0$ and $\theta_{l_2} - \theta_{l_1} \neq 0$, which is in contrast to SCC where both these expressions equal zero). The position of the maximum of the correlation function is dependent on the phase at $u = \pm u_s$ and as the phase at $u = \pm u_s$ is non-zero, due to the difference in window length, the ACC technique incorrectly predicts the shift between $s_1(t)$ and $s_2(t)$ as being $\frac{\theta_{l_1} - \theta_{l_2}}{2\pi u_s}$. When SCC is used, both windows are of the same length and the time delay is computed correctly.

To derive an expression for the phase error which arises from the difference in window lengths, the situation where an auto-correlation is computed using truncated signals of different length is considered for reasons of simplicity and clarity. An expression for the phase error can now be derived by considering the phase of the cross-spectrum of the two signals

$$s_1(t) = s_2(t) = \cos(2\pi u_s t),$$

when each of the signals is truncated by a window function of a different length. For simplicity, let the window used to truncate $s_1(t)$ be an integer number of periods of $s_1(t)$ in length, i.e let $l_1 = pT_s$, where $T_s = 1/u_s$ and p is an integer. Also let $w_2(t)$ have a length l_2 which is a

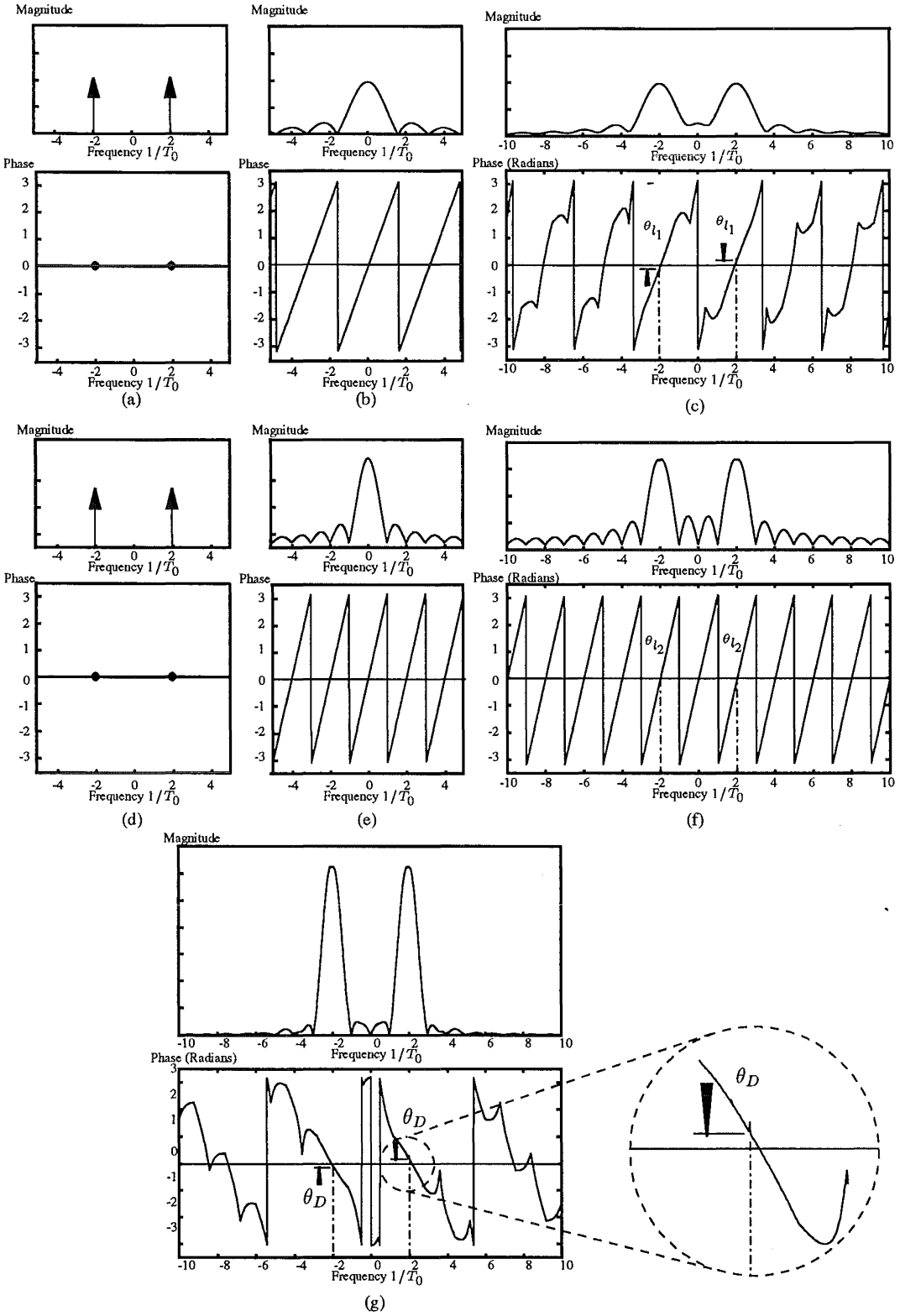

Figure 9.4 The origins of the error that arises when the ACC technique is used to compute a cross-correlation. In the above figure an auto-correlation is shown for simplicity as the ACC technique produces the same errors in the auto-correlation function. In (c) the convolution of the cosine wave's spectrum in (a) and the window function's spectrum in (b) are shown. In (f) the convolution of the cosine wave's spectrum in (d) with the window function's spectrum in (e) are shown. The window function pertaining to (b) has a length 0.625 times that of the window function pertaining to (e). The frequency plots have been normalised with respect to the frequency of the cosine wave. The cosine wave has a frequency $2/T_0$. The different values of the phases of the convolution spectra at $2/T_0$ (θ_{l_1} and θ_{l_2}) can be seen in (c) and (d). The cross-spectrum of (c) and (f) is given in (g) and the non-zero phase at $2/T_0$ which shifts the peak of the cross-correlation can be seen.

fraction of the length of $w_1(t)$, i.e. let $l_2 = apT_s$, where a is a rational number, in the range 0-1, describing the ratio of the window lengths. As the spectrum of a rectangular window of length $2L$ is given by an expression of the form

$$W(u) = \left|\frac{A}{\pi u}\right| |\sin(2\pi Lu)| exp^{-i2\pi Lu}, \tag{9.5}$$

where A is the amplitude of the window (Brigham, 1988), the spectrum of the truncated version of $s_2(t)$ is given by

$$\begin{aligned}(S_2(u) \odot W_2(u)) &= W_2(u+u_s) + W_2(u-u_s) \\ &= \left|\frac{A}{\pi(u+u_s)}\right| |\sin(\pi apT_s(u+u_s))| exp^{-i\pi apT_s(u+u_s)} \\ &\quad + \left|\frac{A}{\pi(u-u_s)}\right| |\sin(\pi apT_s(u-u_s))| exp^{-i\pi apT_s(u-u_s)}. \end{aligned} \tag{9.6}$$

It should be noted at this point that if a similar expression is derived for $s_1(t)$ it is found on evaluating $(S_1(u) \odot W_1(u))$ at $u = \pm u_s$, the resulting summation contains only the term $W_1(0)$, since the terms $W_1(2u_s)$ and $W_1(-2u_s)$ are zero respectively. This is a consequence of window $w_1(t)$ being an integer number of periods of $s_1(t)$ in length. The end result of this property of $(S_1(u) \odot W_1(u))$ is that θ_{l_1} is zero since the spectrum of a cosine wave has zero phase at $u = \pm u_s$. To calculate the phase of Eqn.(9.6) at $u = \pm u_s$ it is necessary to consider the form of Eqn.(9.6) as $u \to \pm u_s$. As $u \to u_s$

$$\left|\frac{A}{\pi(u-u_s)}\right| |\sin(\pi apT_s(u-u_s))| \to \left|\frac{A}{\pi(u-u_s)}\right| (\pi apT_s(u-u_s)) = AapT_s$$

and consequently the expressions,

$$(S_2(u_s) \odot W_2(u_s)) = \left|\frac{A}{\pi}\right| \left[\left|\frac{1}{2u_s}\sin(2\pi ap)\right| exp^{-i2\pi ap} + \pi apT_s\right] \tag{9.7}$$

and

$$(S_2(-u_s) \odot W_2(-u_s)) = \left|\frac{A}{\pi}\right| \left[\pi apT_s + \left|\frac{1}{2u_s}\sin(-2\pi ap)\right| exp^{i2\pi ap}\right] \tag{9.8}$$

are obtained. The phase of the Eqn.(9.7) is given by

$$\theta_{l_2} = Ph(S_2(u_s) \odot W_2(u_s)) = \arctan \frac{|\sin(2\pi ap)| \sin(-2\pi ap)}{|\sin(2\pi ap)| \cos(-2\pi ap) + 2\pi ap} \tag{9.9}$$

and the phase of the Eqn.(9.8) is given by

$$-\theta_{l_2} = Ph(S_2(-u_s) \odot W_2(-u_s)) = \arctan \frac{|\sin(2\pi ap)| \sin(2\pi ap)}{|\sin(2\pi ap)| \cos(2\pi ap) + 2\pi ap}. \tag{9.10}$$

Either of the above two equations can be used to describe the error in the computed time delay, since $|\theta_{l_1} - \theta_{l_2}| = |\theta_{l_2}|$ as $\theta_{l_1} = 0$. The absolute error in the computed time delay ϵ_l is therefore given by the expression

$$\left|\frac{\theta_{l_2}}{2\pi u_s}\right|.$$

To confirm whether Eqn.(9.9) and Eqn.(9.10) can correctly predict ϵ_l for the ACC technique in practice the values predicted by the theoretical expressions have been compared to those obtained in practice. To determine the experimental values the ACC of $s_1(t)$ and $s_2(t)$ was

COMPARISON OF THE ACTUAL AND THEORETICAL ERROR IN THE ADAPTED CROSS-CORRELATION TECHNIQUE AS A FUNCTION OF WINDOW LENGTH RATIO

a	Experimental	Theoretical	Theoretical (% T_0)	Δ (% T_0)
0.063	1.055	0.371	5.9 % of T_0	11.0 % of T_0
0.188	0.172	0.263	4.2 % of T_0	1.0 % of T_0
0.313	0.147	0.145	2.3 % of T_0	0.0 % of T_0
0.438	0.074	0.083	1.3 % of T_0	0.1 % of T_0
0.563	0.074	0.066	1.0 % of T_0	0.1 % of T_0
0.688	0.049	0.061	1.0 % of T_0	0.2 % of T_0
0.813	0.049	0.051	0.8 % of T_0	0.0 % of T_0
0.938	0.049	0.041	0.7 % of T_0	0.1 % of T_0
1.000	0.000	0.000	0.0 % of T_0	0.0 % of T_0

Table 9.1 Δ (% T_0) is the difference between the experimental and theoretical results which have been expressed in radians. The experimental results have been calculated by using Eqn.(9.3) on two in-phase sinusoids. The window length of the second window was adjusted to the values of a specified above. The sample spacing in the experimental situation was 0.4 % of T_0 and as a consequence whenever Δ (% T_0) is less than this value the theoretical and experimental values agree to within the resolution of the experimental computation.

computed, using Eqn.(9.3), for various values of a. In this simulation $s_1(t)$ was sampled so that two periods of the signal were represented by 512 samples and the sampling process therefore gave a Δt of $\frac{2\pi}{256} = 0.0245$ radians. The results from this comparison, presented in Table 9.1, confirm that Eqn.(9.9) and Eqn.(9.10) can indeed correctly predict the ϵ_l that arises in practice when the ACC technique is used to estimate time delays. For most of the values of a tabulated the experimental and theoretical values for ϵ_l agree to within Δt.

The above expressions also demonstrate that as the number of periods of $s(t)$ in the first window, $w_1(t)$, tends to infinity (i.e as $p \rightarrow \infty$) ϵ_l, for a particular value of a, decreases. The decrease in ϵ_l with increasing p is to be expected because as $p \rightarrow \infty$ the magnitude of the first term in Eqn.(9.6) at u_s becomes negligible and, similarly, the magnitude of the second term in Eqn.(9.6) at $-u_s$ also becomes negligible. Thus, as $p \rightarrow \infty$, $\theta_{l_2} \rightarrow 0$ and as a result ϵ_l also tends to zero. It can also be seen from Eqn.(9.9) and Eqn.(9.10) that as $a \rightarrow 1$, $\epsilon_l \rightarrow 0$ for a particular value of the integer p. The reason for this is that the side lobes of the spectrum of the window function move towards the main lobe as a increases. This has an effect on Eqn.(9.6) which is similar to increasing p, and as a consequence θ_{l_2} and $\epsilon_l \rightarrow 0$ as $a \rightarrow 1$. In Fig. 9.5 (a) and (b) graphs of ϵ_l as a function of a are depicted for p equal to two and twenty respectively. These graphs confirm the properties of ϵ_l described above. When the signals being correlated are multispectral, it is believed that the accuracy of the ACC technique is going to be dictated for the most part by the phase error that the significant low frequency components introduce, since the superposition of the window function centered on $\pm u_s$ on that centered at $\mp u_s$ is going to contribute less to $S_2(u_s) \odot W_2(u_s)$ as u increases. As the phase errors corresponding to the low frequency components are relatively large, especially when a is small, it is quite possible that the ϵ_l associated with the ACC technique could be quite large in comparison to that for the SCC technique.

9.2.4 Least Squares Time Delay Estimation

The least squares time delay estimation (LSE) technique computes the time delay between two signals by first calculating the mean squared error between the target signal and a shifted

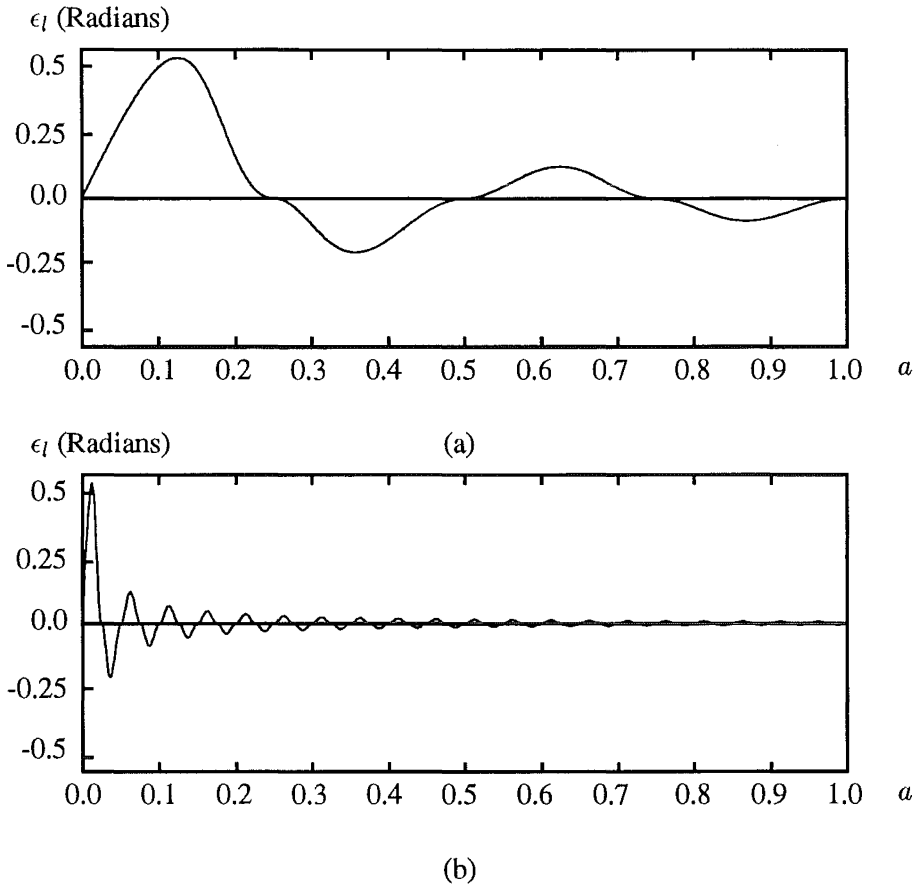

Figure 9.5 The error that arises in the calculated time delay with the ACC technique as a function of a the window length ratio. Shown in (a) is the case where $p = 2$ and in (b) the case where $p = 20$.

version of the response signal (Fertner and Sjölund, 1986), i.e.

$$LSE[i] = \frac{1}{N}\sum_{0}^{N-1}(r[n] - h[n+i])^2, \qquad i = -(N-1) \text{ to } N-1 \tag{9.11}$$

The expectation is that $LSE[i]$ is likely to have its minimum value at the value of i that corresponds to the shift that results in the two signals being overlayed exactly. Thus, by searching for the minimum value of $LSE[i]$ it is possible to determine the time delay between the two signals.

To confirm that the LSE technique can be used to find the time shift between two signals, it is useful to rewrite Eqn.(9.11) using expected value notation. For large N and when $h[n] = r[n+n_\tau] + c[n]$, where c is zero-mean Gaussian-distributed noise, the expression for the mean square error can be expressed using expected value notation in the form

$$\begin{aligned} LSE[i] &= \frac{1}{N}\sum_{0}^{N-1}(r[n] - h[n+i])(r[n] - h[n+i]), \qquad i = -(N-1) \text{ to } N-1 \\ &= \frac{1}{N}\sum_{0}^{N-1}(r[n]^2 - 2r[n]h[n+i] + h[n+i]^2) \\ &= \frac{1}{N}\sum_{0}^{N-1}(r[n]^2 - 2r[n](r[n+n_\tau+i] + c[n+i]) + (r[n+n_\tau+i] + c[n+i])^2) \end{aligned}$$

$$= \frac{1}{N}\sum_{0}^{N-1}(r[n]^2 - 2r[n]r[n+n_\tau+i] - 2r[n]c[n+i] + r[n+n_\tau+i]^2$$

$$+2r[n+n_\tau+i]c[n+i] + c[n+i]^2) \quad (9.12)$$

$$= E\{r[n]^2 - 2r[n]r[n+n_\tau+i] + r[n+n_\tau+i]^2 + c[n+i]^2\} \quad (9.13)$$

The $E\{c[n+i]^2\}$ term is the variance of the noise and the $E\{r[n]^2\}$ and $E\{r[n+n_\tau+i]^2\}$ terms are both the mean square value of the signal, which in the situation where $E\{r[n]\} = 0$ are equivalent to the variance of the signal. For stationary ergodic data and for large N these terms should be constant, hence the amount the response signal lags behind the target signal is then found by determining for which value of i $LSE[i]$ is a minimum. On comparing Eqn.(9.13) to Eqn.(9.2) it can be seen that when N is large and the signals and contamination are stationary the SCC and LSE techniques are equivalent. A minor advantage of the LSE technique is, however, that the dc component in each signal does not have to be removed prior to computing the expression in Eqn.(9.11) as the subtraction in Eqn.(9.11) should inherently remove the dc component. It can also be seen that if the summation in Eqn.(9.11) is only calculated over those values of the signal which are non-zero (i.e. the contribution from the overlapping sections of the signal which are zero is ignored), in the uncontaminated situation, the LSE technique is able to estimate the time delay exactly unlike either the SCC or ACC techniques. In the study described in §9.2.6 this technique was used.

9.2.5 Phase Shift Time Delay Estimation

9.2.5.1 Description of Technique

The phase shift time delay estimation technique (PSE) (Hunter *et al.*, 1987; Brigham, 1988; Rababy *et al.*, 1989; Wu *et al.*, 1989) uses the Fourier transform relationship

$$f[n-n_\tau] \overset{FT}{\Longleftrightarrow} F[u]\exp^{i2\pi u n_\tau}, \quad (9.14)$$

which states that if a signal is delayed by a time n_τ then the phase of its spectrum undergoes a linear phase shift $\exp^{i2\pi u n_\tau}$ (Table 2.1). To compute the time delay, the phase difference between each of the spectral components of the response and target signals is first calculated. These phase differences are then used to find the slope b of the linear equation $p[u] = bu$, which describes the linear phase difference between the two signals. To find the slope, a least squares linear approximation technique is used to find the slope of the line which best fits the obtained phase difference data (Burden and Faires, 1985). In the uncontaminated situation, the slope b computed by this technique is equal to $2\pi n_\tau$ and it is therefore possible to calculate the time delay n_τ. Thus, this time delay estimation technique inherently computes time shifts to sub-sample accuracy unlike the other techniques which compute n_τ to the nearest multiple of Δt.

In practice, the response signal is usually contaminated as well as being a time shifted version of the target signal and, as a consequence, weighted least squares techniques are usually used to compute b. The contamination is less likely to significantly perturb the phases of those frequency components of $F[u]$ which have the largest magnitude and so the weights usually correspond to the magnitudes of the frequency components of one of the signals. The author used this approach in the study investigating the accuracy of the five time delay estimation techniques. A more sophisticated approach is to use the coherence function to weight the data values in the least squares technique (Hunter *et al.*, 1987; Rababy *et al.*, 1989).

The coherence is a measure of the degree of correlation between two signals as a function of frequency (Bendat and Pierslo, 1980; Hunter *et al.*, 1987). When two signals have frequency components that are well correlated the coherence function has a value approaching one and when the frequency components are uncorrelated the coherence function has a value of zero. Thus, using the coherence function, in conjunction with the weighted least squares technique, allows those data points corresponding to the signals' well correlated frequency components to be weighted more heavily than those components where the correlation is poor.

9.2.5.2 Non-Ideal Characteristics of the Technique

To determine what effect contamination on the response signal has on the computed phase differences, consider the response signal $h[n] = f[n - n_\tau] + c[n]$. On invoking the Fourier transform the spectrum of the response signal,

$$H[u] = F[u]\exp^{i2\pi u n_\tau} + C[u],$$

is obtained. This expression can be rewritten in the form

$$|H[u]|\exp^{iPh(H[u])} = |F[u]|\exp^{iPh(F[u])}\exp^{i2\pi u n_\tau} + |C[u]|\exp^{iPh(C[u])}.$$

The phase difference, $Ph_{fh}[u]$, between the two signals $f[n]$ and $h[n]$ can then be described by the expression

$$\begin{aligned} Ph_{fh}[u] &= Ph(H[u]F^*[u]) \\ &= Ph\left(\left[|F[u]|\exp^{iPh(F[u])+2\pi u n_\tau} + |C[u]|\exp^{iPh(C[u])}\right]|F[u]|\exp^{-iPh(F[u])}\right) \\ &= Ph\left(\exp^{i2\pi u n_\tau} + \frac{|C[u]|}{|F[u]|}\exp^{iPh(C[u])-Ph(F[u])}\right) \end{aligned} \tag{9.15}$$

It can be seen that the vector describing the required phase difference is perturbed by another vector which has a magnitude dependent on the SNR of $h[n]$ and a phase which is the difference between the phase of the contamination and the original signal at a particular frequency. Thus, as the level of contamination rises and the SNR decreases, the linear phase relationship between the two signal's spectra becomes increasingly distorted. In the situation where the contamination has a flat spectrum, the distortion at each value of u is random, because the phase of $C[u]$ at a particular value of u is completely random. The effect of the contamination is to therefore randomly distribute the $Ph_{fh}[u]$ about $2\pi n_\tau u$.

As the time delay is computed solely in the frequency domain, the time delay estimate is especially susceptible to contamination introduced by the incorrect application of the discrete Fourier transform. The two undesirable artifacts resulting from use of the DFT are aliasing and spectral leakage (see §2.5). As explained in §2.5 aliasing occurs when a signal is not sampled at a rate at least as great as twice its highest frequency and spectral leakage occurs because in practice all signals must be truncated before they are Fourier transformed. The effect of both of these artifacts is to distort the spectra of the two signals being transformed and to randomly distribute the $Ph_{fh}[u]$ about $2\pi n_\tau u$, making accurate estimation of the time delay difficult. If the bandwidth of $F(u)$ is known *a priori* it is possible to eliminate the effect of aliasing by setting the sample spacing so that the signals are sampled at twice their original bandwidth and $H(u)$ can be pre-filtered to remove those components of $C(u)$ that are above the original bandwidth. In the author's study a four-pole low pass digital Butterworth filter

(see Ludeman (1987)) was used to remove those components of $C(u)$ that were above 0.2 Hz. The effect of spectral leakage cannot be eliminated entirely, but can be reduced by windowing the signals with a window function that has more compact sidelobes. In the author's study the Hanning and Blackman windows were used in an attempt to reduce the distortion caused by spectral leakage.

9.2.6 The Accuracy of Time Delay Estimation Techniques: An Experimental Comparison by Computer Simulation

9.2.6.1 Methods

The ability of the standard spatial domain cross-correlation, adapted cross-correlation, least squares time delay estimation, Fourier domain cross-correlation, and phase shift delay estimation techniques to accurately estimate the time delay that exists between a tracking task's target signal and a subject's response signal has been compared by computer simulation. To determine which method could provide the most accurate time delay estimate, the absolute difference between the simulated and computed time delay, denoted hereafter by ϵ_l, was computed for simulated signals with various SNRs and time delays. The mean ϵ_l and standard deviation (s.d.) of ϵ_l were calculated for various noise levels so the robustness of the various techniques, with regards to the noise level, could be determined.

The target signal used in this comparison was that described in §8.4, that is, a waveform constructed from 28 equal amplitude harmonically related sinewaves with a fundamental frequency of 0.007 Hz and randomly selected phases, which is an approximation to 0.2 Hz bandlimited white noise. The target waveform is represented as 4096 samples which are separated by a time interval (Δt) of 35 ms (i.e. a sample rate of 28.57 Hz). In tracking situations where the subject's S-M system is not being unduly stressed the subject's response to the target signal $x[n]$ can be modelled by the expression

$$y[n] = x[n + n_\tau] + c[n], \tag{9.16}$$

where $y[n]$ is the signal describing the subject's response, $n_\tau \Delta t$ is the time delay between the target and response signals, and $c[n]$ is a term which models the error in the subject's tracking (Lynn *et al.*, 1979; Trier Andersen, 1986). Here, $c[n]$ is modelled by zero mean Gaussian distributed noise. In practice it seems very likely that the error could be signal dependent, however all studies to date have modelled the error with zero-mean Gaussian distributed noise (see §9.2.1).

The resulting set of simulated response signals had SNRs of ∞, 40 dB, 30 dB, 20 dB or 10 dB. To simulate time delays experienced in practice, each of these five noisy signals were then shifted by an integer number of samples varying between zero and fifteen (i.e. delays ranging from 0 to 525 ms), producing seventy five different signals in total. The values used here for the time delays and SNRs are considered to cover the range of time delays and levels of noise that most likely characterise the responses of normal subjects and most subjects with mild S-M disorders. In the study of normal and Parkinsonian S-M function described in §10.1 it was found that the subject's produced time delays in the range zero to twelve samples and response signals with SNRs of 12.3 dB to 33.3 dB.

For each of the seventy five signals, two 1024 sample sections of the complete signal (beginning at the 30th and 60th sample of the complete signal respectively) were presented to the different time delay estimation techniques. It should be noted that within these 1024

sample sections of the response signal there is approximately one quarter of a period of the fundamental and approximately seven periods of the highest frequency component. The ACC technique therefore has a value for p for the fundamental frequency component of 0.25 and the window length ratio, a, is 0.5. To evaluate whether the accuracy of the techniques was sensitive to the section of the response signal used to compute the time delay, a set of signals formed by extracting different 1024 sample sections from the noiseless response signals were also given to the techniques. This set of signals had simulated time delays of 0, 5, 10, and 15 samples and were constructed from 1024 sample sections of the complete simulated response signal beginning at the 90th, 130th, 180th, 230th, 300th, and 360th samples.

9.2.6.2 Results

Table 9.2 shows how the mean time delay estimate errors of the various techniques' change as the SNR of the simulated response signal changes. At each SNR, the time delay estimation errors obtained from the thirty different simulated response signals (i.e the signals with time delays of zero to fifteen samples obtained from the two different sections of the complete response signal) were averaged to give the mean ϵ_l, which is denoted here by $\overline{\epsilon_l}$. The results in Table 9.2 clearly show that the SCC and LSE techniques are the most accurate time delay estimation techniques for the range of SNRs examined. At a SNR of 10 dB the $\overline{\epsilon_l}$ produced by the SCC technique was less than one sample and for the LSE technique the $\overline{\epsilon_l}$ was approximately equivalent to two samples. This is in agreement with the studies reviewed in §9.2.1 which found that in most cases where the SCC technique was compared to other time delay estimation techniques the SCC technique tended to provide the most accurate estimates of the time delay. It can be seen in Table 9.2 that the results obtained from the SCC technique are indeed equivalent to those obtained from the FCC technique and therefore the above statements, and those which follow, concerning the SCC technique apply equally well to the FCC technique. It needs to be noted that the dc components of the signals were removed prior to computing the time delay by SCC and as a consequence, any possibility of the dc components biasing the time delay estimates, as described in §9.2.2, has been removed.

COMPARISON OF THE ACCURACY OF FIVE TIME DELAY ESTIMATION TECHNIQUES FOR VARIOUS SNRs

SNR (dB)	SCC $\overline{\epsilon_l}$ (ms)	FCC $\overline{\epsilon_l}$ (ms)	ACC $\overline{\epsilon_l}(ms)$	LSE $\overline{\epsilon_l}$ (ms)	PSEH $\overline{\epsilon_l}$ (ms)	PSEB $\overline{\epsilon_l}$ (ms)
∞	0.0	0.0	69.8	0.0	3.0	2.1
40	0.0	0.0	69.8	0.0	1.8	2.7
30	0.0	0.0	87.2	0.0	5.6	4.0
20	0.0	0.0	122.1	0.0	48.1	48.3
10	17.5	17.5	314.1	69.8	256.6	224.5

Table 9.2 The above results should be compared to the sample spacing of 35 ms. $\overline{\epsilon_l}$ is the average of the time delay estimation errors, obtained from signals of the same SNR with delays of 0 to 15 samples and for the response signals extracted from various sections of the complete signal (see §9.2.6.1). PSEH and PSEB are the results from the PSE technique with the Hanning and Blackman window respectively.

The time delay estimation inaccuracy of the ACC technique which motivated this study can also be clearly seen in Table 9.2. When the SNR of the simulated response signal was 10 dB the ACC technique had a $\overline{\epsilon_l}$ of a massive nine samples (compared with less than one sample for the SCC technique) and in the uncontaminated situation had an equally disturbing error of approximately two samples. It is shown latter that even in the uncontaminated situation the ACC technique can have a $\overline{\epsilon_l}$ of significantly more than two samples. The reader should not

try to reconcile, for the ACC technique, the values in Table 9.2 for $\overline{\epsilon_l}$ to those in Table 9.1 since Table 9.1 was produced for a monochromatic signal with $p = 2$. The results from the PSE technique show that this technique works reasonably well up to SNR $\geq$ 30 dB after which ϵ_l begins to grow rapidly to give a mean ϵ_l of approximately seven samples at a SNR of 10 dB. The target and simulated response signals had only fifteen significant frequency components within their bandwidth and this was an insufficient number of frequency components for the weighted least squares technique to accurately estimate the phase difference in the contaminated situation. For signals of a larger bandwidth it is likely that the performance of the PSE technique would improve for the range of SNRs in Table 9.2. The results in Table 9.2 also indicate that with the PSE technique there does not appear to be any advantage in using the Blackman window in preference to the Hanning Window. It is also interesting to note that with all the techniques, except for the PSE technique at low SNRs, ϵ_l did not vary significantly as the time delay between the two signals was changed. This is to be expected as a shift of up to fifteen samples is negligible compared to the 1024 samples used to represent the signals and, therefore, the statistics of the sections of signal used to compute the delay do not change appreciably as the time delay was changed.

The signals used here to compute the time delays were extracted from much larger signals and the sections of the extracted signal contained approximately one quarter of a period of the fundamental. As a result of spectral leakage and aliasing, the spectra of the signals varied quite markedly as the section extracted was varied. To determine how susceptible the accuracy of each of the time delay estimation techniques was to the variation in the section of signal extracted to compute the time delay a study, using uncontaminated signals extracted from different regions of the complete signal, was performed (see §9.2.6.1). The results from this study are presented in Table 9.3 where the superiority of the LSE technique in the uncontaminated situation can be clearly seen. As described in §9.2.1 and §9.2.4, in the uncontaminated situation, the LSE technique is able to obtain an exact estimate of the time delay, provided that the time delay is not too large in comparison to the length of the signals being used to compute the delay. In the uncontaminated situation the PSE technique also works well as the slope estimation technique can accurately compute the slope of the linear equation describing the delay as the only perturbations on the data are the small errors due to the finite number representation scheme in the computer used, and spectral leakage and aliasing due to the truncation of the signal. There is is a small variation in ϵ_l (i.e. the s.d. of ϵ_l is 2.5 ms) which is probably due to the aliasing and spectral leakage perturbing the weights (i.e. the signal's spectral magnitude) used in the slope computation.

A small inherent bias of approximately one sample in the time delay estimates computed by the SCC technique, which could be due to the effective multiplication of the cross-correlation function by a triangular window function (see §9.2.2.2), can be seen in Table 9.3. There is a small variation in ϵ_l (i.e. the s.d. of ϵ_l is approximately one sample) as the section of extracted signal is varied and this is most probably due to the changes which aliasing and spectral leakage induce in the high frequency components of the cross-spectrum. The worst result in Table 9.3 is that for the ACC technique which has a mean ϵ_l of approximately seven samples for the uncontaminated case where the sections of extracted signal were varied. This large difference between the mean ϵ_ls obtained from the SCC and ACC (i.e. 1 sample *vs* seven samples) confirms the belief that the inherent bias with the ACC technique is likely to be much worse than that associated with the SCC technique. The significant size of the ϵ_l associated with the ACC technique is largely due to the small window length ratio ($a = 0.5$) and the small values of p for the low frequency components (e.g. p for the fundamental was 0.25) producing large phase errors in the cross-spectrum (see §9.2.3.2). If a was close to unity

COMPARISON OF THE ACCURACY OF FIVE TIME DELAY ESTIMATION TECHNIQUES FOR VARIOUS SECTIONS OF UNCONTAMINATED SIGNAL

Technique	$\overline{\epsilon_l}$ (ms)	s.d. ϵ_l (ms)
SCC	31.0	40.7
FCC	38.8	53.6
ACC	197.7	219.3
LSE	0.0	0.0
PSEH	2.6	2.5
PSEB	2.4	2.5

Table 9.3 The above results should be compared to the sample spacing of 34.9 ms. $\overline{\epsilon_l}$ is the average of the time delay estimation errors, obtained from uncontaminated signals with delays of 0, 5, 10 and 15 samples and for the response signals extracted from various sections of the complete signal (see §9.2.6.1). PSEH and PSEB are the results from the PSE technique with the Hanning and Blackman window respectively. s.d. is the standard deviation of the ϵ_l

and the values of p corresponding to the low frequency components were greater than unity, the ϵ_l associated with the ACC technique could possibly approach that of the SCC technique. The aliasing and spectral leakage introduces phase changes in the low frequency components of the cross-spectrum and this is the most likely reason for the considerable variability in ϵ_l as the section of signal extracted to compute the time delay by ACC was varied.

In summary, the results above have demonstrated that the belief that the ACC technique is likely to provide inaccurate estimates of the time delay between two signals was justified. The study indicated that the most accurate time delay estimation techniques, which can be used in the tracking environment, are the SCC, FCC and LSE techniques. As a result of this study the author chose to use the FCC technique to compute the time delays for the 2-D tracking task described in §9.1.1.

CHAPTER 10

S-M PERFORMANCE OF NORMAL AND PARKINSONIAN SUBJECTS IN 2-D

Once the 2-D tracking tasks described in §9.1 had been completed the author used them in two studies designed to investigate various aspects of normal and abnormal S-M function. The primary objectives of the first study were to verify whether the 2-D tasks could demonstrate the S-M deficits known to occur with a common neurological disorder (namely Parkinson's disease), to determine whether there is any difference between abnormal and normal 1-D and 2-D tracking performance, and to determine whether there is any benefit to be gained from using 2-D tracking tasks, as opposed to 1-D tracking tasks, when investigating both normal and abnormal S-M function. The last objective was the most important because, as outlined in §8.3.4, no-one to date has attempted to justify the use of 2-D tracking tasks on the basis that they may provide new, more sensitive, additional, or more valid information on S-M function than 1-D tasks. Normal S-M function in both 1-D and 2-D environments was the subject of the second study and the major objective was to answer several questions regarding the benefits of 2-D tasks that were not clearly answered by the first study.

In this chapter the two studies which have been used to investigate both normal and Parkinsonian S-M function in 1-D and 2-D are presented. These studies represent a significant portion of the author's original research in the S-M area. In §10.1 the pilot study investigating normal and Parkinsonian S-M function in multidimensional environments is described. This is followed in §10.2 by the study of normal 1-D and 2-D tracking performance which was designed to provide a definitive answer to the question of whether tracking in 2-D is more difficult than tracking in 1-D. In the respective sections the objectives, experimental procedure, results, and discussion of the results from each of the studies is presented.

10.1 STUDY ONE: STUDY COMPARING NORMAL AND PARKINSONIAN S-M FUNCTION IN 2-D

10.1.1 Objectives

An experimental study of normal and Parkinsonian S-M function was conducted using the 2-D tracking tasks described in §9.1. The motivation for conducting the study was the need to verify the validity of the 2-D tracking tasks and to verify the beliefs, outlined in §8.3.4 regarding the benefits to be gained from using 2-D tracking tasks in preference to 1-D tracking tasks. More specifically, the aims of the study were to address the following questions:

- Do 2-D tracking tasks demonstrate deficits in Parkinson's diseases similar to those reported for 1-D tracking?
- Do either or both groups find 2-D tracking more difficult than 1-D tracking? If there is a difference, is it greater for Parkinsonian subjects?
- Do either or both groups find tracking targets moving in a horizontal direction easier than in the vertical direction, or vice versa?
- To what extent can each group make use of spatial predictability to improve their performance?
- Can 2-D tracking tasks provide more sensitive measures than 1-D tracking tasks of, or new information on, sensory-motor function?

10.1.2 Experimental Design

10.1.2.1 Subjects

Eight subjects with idiopathic Parkinson's disease (PD) (two female and six male) and eight normal control subjects (six female and two male) participated in the study. The Parkinsonian subjects were mildly affected (I and II on the Hoehn-Yahr Scale (Hoehn and Yahr, 1967)) and remained on their medication during the course of the study. All subjects were screened for major head injury, drug and alcohol problems, and previous psychiatric or non-Parkinsonian neurological disorders which could affect the outcome of the study. The Parkinsonian and normal groups were matched with respect to age (65.7 *vs* 63.7 years, $p>0.1$) and level of education (10.0 *vs* 10.2 years, $p>0.1$). The probability (p) values indicate that there is no significant difference between the ages or level of education of the two groups.

10.1.2.2 Experimental Setup

Both the 2-D and 1-D versions of the 2-D tracking tasks were used in the study with the floor-mounted joystick as the input device in all cases. The subjects were seated in front of the stimulus display monitor, with an eye-to-monitor distance of approximately 0.9 m. The joystick was placed by the front leg of the subject's chair and had the hand-rest fixed at a height of 0.61 m, resulting in the subject's forearm being horizontal and the elbow bent with an angle of approximately 90 deg. The task duration, target size, and target signal used for the random pursuit tasks in the study were those specified by the default settings (see §9.1).

10.1.2.3 Experimental Procedure

The sixteen subjects were divided into four subgroups, each containing two subjects from the Parkinsonian group and their matched controls from the normal group. All subjects completed the 2-D random, 2-D step, and 1-D horizontal and vertical random tracking tasks with the order such that any order effects would be balanced out; that is, half the subjects completed a 1-D tracking task before a 2-D random task and half completed a 1-D vertical tracking task before a 1-D horizontal task.

The subjects carried out the tasks with the hand which felt most comfortable, resulting in three Parkinsonian subjects completing the task with a hand which did not correspond to the

DESCRIPTION OF EXPERIMENTAL AND CONTROL SUBJECTS

Subject	Age (years)	Gender	Education (years)	Dominant Hand	Hand Used	Affected Side
			Parkinsonian Subjects			
1	51	M	11	R	R	R
2	67	M	9	AM	R	B
3	77	F	9	R	R	L
4	57	M	13	AM	L	R
5	70	F	9	R	R	R
6	64	M	9	R	R	B
7	73	M	11	R	R	B
8	67	M	9	R	R	L
			Normal Subjects			
1	51	F	11	R	R	N/A
2	61	F	9	R	R	N/A
3	63	F	9	R	R	N/A
4	62	M	11	R	R	N/A
5	70	F	8.5	R	R	N/A
6	65	F	10.5	R	R	N/A
7	70	M	15	L	L	N/A
8	68	F	7.5	R	R	N/A

Table 10.1 R=Right, L=Left, AM=Ambidextrous and B=Both.

laterality of their disorder (see Table 10.1). It is important to note that all subjects had had some previous exposure to the 2-D random tasks earlier in the same session as part of an independent parallel study of 'central executive' deficits in Parkinson's Disease (Dalrymple Alford *et al.*, 1994).

10.1.2.4 Analysis

The BMDP statistical package (Dixon, 1981) was used to perform Wilcoxon matched pairs comparisons of the results in this study. Non-parametric statistics were considered to be most appropriate as tracking performance distributions tend to be highly skewed and have quite different variances. Statistical tests were one-tailed for inter-group comparisons and two-tailed for intra-group comparisons. In the results that are presented in the following section the difference in the two group's means is expressed as a percentage ($\Delta\%$) of the mean of either the normal group's performance or the performance observed on the 1-D task. When this difference is observed for a significant number of the matched pairs comprising each of the distributions this is indicated by giving the probability (p) value. Any test for which $p > 0.1$ is marked not significant (NS) which indicates that any signed difference between the performance of the two groups was not observed for a significant number of the matched pairs.

The nomenclature for the measures on the 1-D tasks is the same as that for the 2-D tasks except 2D is replaced by 1D (i.e. $MAE2D_X$ becomes $MAE1D_X$). Definitions of the nomenclature in terms of the performance measures can be found in §9.1. It is necessary to note that the 1-D step task performance measures used in the majority of the analyses were the mean values of the five minimum reaction times, rise times, overshoots, target entry times, and

errors. As described §9.1.2 during the course of this study it was found that the above condensed performance measures were unsuitable and for this reason they were replaced with the median performance measures described in §9.1.2 for some analyses. Where the median performance measures have been used to obtain the results that follow, this is explicitly indicated.

10.1.3 Results

Do 2-D tracking tasks demonstrate deficits in Parkinson's diseases similar to those reported for 1-D tracking? The random tracking results confirm that Parkinsonian subjects cannot track a random signal as accurately as normal subjects. The Parkinsonian group had larger absolute errors ($MAE2D_R$: 11.5 mm *vs* 4.7 mm, 145%, p<0.01; $MAE2D_X$: 8.5 mm *vs* 3.3 mm, 156%, p<0.05; $MAE2D_Y$: 7.6 mm *vs* 3.3 mm, 131%, p<0.01; $MAE1D_X$: 6.1 mm *vs* 3.0 mm, 104%, p<0.01; $MAE1D_Y$: 6.4 mm *vs* 3.1 mm, 107%, p<0.05) and larger lags ($Lag2D_X$: 221 ms *vs* 119 ms, 86%, p<0.05; $Lag2D_Y$: 286 ms *vs* 154 ms, 86%, p<0.1; $Lag1D_X$: 156 ms *vs* 84 ms, 86%, p<0.1; $Lag1D_Y$: 151 ms *vs* 85 ms, 78%, NS). There was no significant difference between the gains for the two groups.

For movements to the outside target squares on the step task the Parkinsonian group's reaction times (678 ms *vs* 547 ms, 24%, p<0.05), target entry times (2548 ms *vs* 1690 ms, 51%, p<0.01) and overall errors (23% *vs* 17%, 35%, p<0.05) were all greater than those for the normal group. The Parkinsonian group had a small negative mean overshoot (-1.15%), whereas the normal group had a small positive mean overshoot (1.05%). On looking at the step response graphs for the Parkinsonian group, the classic Parkinsonian step response of a stepped approximation to an overdamped response (Flowers, 1976; Day *et al.*, 1984) was observed in the majority of cases. Nonetheless, several Parkinsonian subjects produced multiple underdamped responses.

Do either or both groups find 2-D tracking more difficult than 1-D tracking? If there is a difference, is it greater for Parkinsonian subjects? Comparison of overall error measures showed that the 2-D performance of both normal and Parkinsonian subjects was inferior to their 1-D performance (Norm $MAE2D_R$ *vs* $MAE1D_X$: 4.7 mm *vs* 3.0 mm, 57%, p<0.01; Norm $MAE2D_R$ *vs* $MAE1D_Y$: 4.7 mm *vs* 3.1 mm, 52%, p<0.01; PD $MAE2D_R$ *vs* $MAE1D_X$: 11.5 mm *vs* 6.1 mm, 87%, p<0.01; PD $MAE2D_R$ *vs* $MAE1D_Y$: 11.5 mm *vs* 6.4 mm, 79%, p<0.01). A significant increase in the lag, in one coordinate, was observed for both normal and Parkinsonian subjects when the 2-D task lags were compared to the 1-D task lags. For the normals the most significant increase occurred with the X-coordinate (119 ms *vs* 84 ms, 42%, p<0.1), whilst that for the Y-coordinate was not significant (153 ms *vs* 85 ms, 82%, NS). The Parkinsonian subjects had a significant increase in the lag in the Y-coordinate (286 ms *vs* 151 ms, 90%, p<0.01) but not in the X-coordinate (221 ms *vs* 156 ms, 42%, NS). On comparing the difference in performance between 1-D and 2-D tasks in normals as opposed to Parkinsonian subjects, it was found, for both the additive and proportional models of performance degradation, that the Parkinsonian subjects' performance deteriorated significantly more than the normals' (Proportional: 40% *vs* 29%, p<0.1; Additive: 5.2 mm *vs* 1.6 mm, p<0.01).

As reported in §8.3.4, Garvey and Taylor (1959), Chernikoff *et al.* (1960), Todosiev *et al.* (1967), and Ziegler (1968) evaluated the effect of adding the extra dimension in a 2-D task by comparing the $MAE1D_X$ and $MAE1D_Y$ to the $MAE2D_X$ and $MAE2D_Y$ respectively to see if there was any degradation in performance. In this study comparing the 1-D performance measures to those from the same coordinate in the 2-D task showed that for the normals there

was no difference in performance (X: 3.3 mm *vs* 3.0 mm, 10 %, NS; Y: 3.3 mm *vs* 3.1 mm, 6 %, NS) while for the Parkinsonian subjects there was a difference in performance (X: 6.1 mm *vs* 8.5 mm, 39 %, $p<0.1$; Y: 6.4 mm *vs* 7.6 mm, 19 %, $p<0.05$)

Do either or both groups find tracking targets moving in a horizontal direction easier than in the vertical direction, or vice versa? No directional bias was found on the random tracking for either the normal group (2-D: 3.3 mm *vs* 3.3 mm, 1%, NS; 1-D: 3.0 mm *vs* 3.1 mm, -3%, NS) or the Parkinsonian group (2-D: 7.6 mm *vs* 8.5 mm, 12%, NS; 1-D: 6.1 mm *vs* 6.4 mm, -4%, NS). The same was also true for the equivalent lags and gains. Similarly, no differences were found, for Parkinsonian or normal subjects, between the horizontal and vertical components of the step pursuit tracking task's performance measures (see Table 10.2), with the notable exception of the target entry times. It should be noted that the step task performance measures used in this comparison were the median values of the performance measures as described in §9.1.2. This allowed each performance measure to be directly compared to each other since each has now been calculated from the same targets. A comparison of the target entry times for both the normal group (outside targets: 2240 ms *vs* 1645 ms, 36%, $p<0.01$; central target: 1524 ms *vs* 1350 ms, 13% NS) and the Parkinsonian group (outside targets: 2871 ms *vs* 1955 ms, 47%, $p<0.01$; central target 2440 ms *vs* 1642 ms, 49%, $p<0.05$) demonstrated that performance in the vertical direction was superior to that in the horizontal direction for both groups.

Further analysis relating to this result confirmed that, for normal subjects, the distributions of overshoots were near identical in the horizontal and vertical directions with standard deviations of 7.99% and 8.85% respectively. Together with similar reaction times, risetimes, and overshoots in both directions, this indicates that the responses in each direction were similar up until the first overshoot and discounts the possibility that similar primary movement measures resulted from different distributions of responses in the two directions. Inspection of graphs of the normal subjects' responses indicated a tendency to adopt different tracking strategies in the horizontal and vertical directions after the first overshoot. During this secondary correction phase, normal subjects tended to exhibit an overdamped tracking strategy in the vertical direction whilst maintaining underdamped characteristics in the horizontal direction. Consequently the subjects' responses tended to enter the target earlier in the vertical direction, as evidenced by its lower target entry time. For the Parkinsonian subjects the overshoot distributions were different in horizontal and vertical directions (e.g. standard deviations of 17.45% and 9.59%). It is therefore not possible to draw similar conclusions regarding the independence of secondary corrections from primary movements in the Parkinsonian responses.

VERTICAL AND HORIZONTAL PERFORMANCE ON THE STEP TRACKING TASK.

	Parkinsonian Subjects				**Normal Subjects**			
	Outside Targets		Central Targets		Outside Targets		Central Target	
Error Measure	Δ%	p	Δ%	p	Δ%	p	Δ%	p
Reaction Time	-8	NS	4	NS	-1	NS	2	NS
Rise Time	-12	NS	-7	NS	0	NS	-4	NS
Overshoot	41	NS	11	NS	-61	NS	46	NS
Target Entry	47	<0.01	49	<0.05	36	<0.01	13	NS
Error	-4	NS	0	NS	1	NS	3	NS

Table 10.2 A negative Δ% indicates that the horizontal performance is superior to the vertical performance.

To what extent can each group make use of spatial predictability to improve their

performance? Both the normal and Parkinsonian subjects were able to use the spatial predictability of the central target to improve their tracking performance. This resulted in a reduction in errors (Norm: 15% *vs* 17%, -11%, $p<0.05$; PD: 18% *vs* 23%, -19%, $p<0.01$), reaction times (Norm: 481 ms *vs* 547 ms, -12%, $p<0.01$; PD: 457 ms *vs* 678 ms, -48%, $p<0.01$), and target entry times (Norm: 1306 ms *vs* 1689 ms, -23%, $p<0.05$; PD: 2027 ms *vs* 2548 ms, -20%, $p<0.01$) for both groups. As expected, there were no changes in rise times. In addition, there was no difference between the two groups in the level of improvement, whether for the additive model of performance improvement (reaction time: 221 ms *vs* 66 ms, NS; rise time: -31 ms *vs* 38 ms, NS; target entry: 521 ms *vs* 383 ms, NS; error: 4% *vs* 2%, $p<0.1$) or the proportional model (reaction time: 30% *vs* 12%, NS; rise time: -8% *vs* 12%, NS; target entry: 23% *vs* 21%, NS; error: 19% *vs* 11%, NS;).

Can 2-D tracking tasks provide more sensitive measures than 1-D tracking tasks of, or new information on, sensory-motor function? Differences in the mean absolute errors between the two groups for the 2-D task ($MAE2D_R$: 11.5 mm *vs* 4.7 mm, 145%, $p<0.01$; $MAE2D_X$: 8.5 mm *vs* 3.3 mm, 156%,$p<0.05$; $MAE2D_Y$: 7.6 mm *vs* 3.3 mm, 131%, $p<0.01$) were all greater than those for the 1-D tasks ($MAE1D_X$: 6.1 mm *vs* 3.0 mm, 104%, $p<0.01$; $MAE1D_Y$: 6.4 mm *vs* 3.1 mm, 107%, $p<0.05$). In addition, there was a greater deterioration in Parkinsonian performance on going from a 1-D to a 2-D task than there was for the normal subjects (proportional model: 40% *vs* 29%, $p<0.1$; additive model: 5.2 mm *vs* 1.6 mm, $p<0.01$). These results demonstrate that the 2-D task was more sensitive, although not markedly so, than the 1-D tasks in distinguishing normal and impaired sensory-motor performance.

10.1.4 Discussion

A study of 1-D and 2-D tracking performance in both normal and Parkinsonian subjects has been conducted with the 2-D tasks and their 1-D counterparts. Results from this study confirm that the 2-D tracking tasks can detect and quantify the sensory-motor deficits expected in Parkinson's disease. These deficits were revealed by a degradation in overall tracking performance on the random and step tasks. In addition, with the 2-D step task it was noted that many Parkinsonian subjects' responses were the classic stepped approximation to an overdamped response, which confirms the inability of the Parkinsonian subjects to effectively perform open-loop movements (Flowers, 1976; Day *et al.*, 1984), while the normal subjects generally adopted underdamped open-loop response strategies. It is unlikely that these deficits could be attributed to the difference in gender between the two groups since any gender bias (should it exist) would tend to favour the Parkinsonian group, as the tracking performance of males tends to be superior to females (Jones *et al.*, 1986). The differences in performance observed between the Parkinsonian and normal groups were greater for the 2-D task than for the 1-D tasks. Thus, for neurological disorders in general, the 2-D task should provide a measure of sensory-motor dysfunction which is at least as sensitive, and will often be more so, than its 1-D counterpart. Conceptually, it is not surprising that the sensitivity of the 2-D task should be greater or equal to that of the 1-D tasks, since the 2-D task places additional demands on the subjects which may (or may not) be met by either the normal or impaired sensory-motor system. Further experimental work would be required to determine whether the increased separation between the two groups performance distributions reflects an abnormal adjustment by the Parkinsonian subjects to the extra dimension in the 2-D task or simply reflects the increased speed/acceleration demands in the 2-D task from either a motor and/or visuospatial point of view.

Using overall measures of performance it was noted that there was a significant deterioration in both normal and Parkinsonian tracking performance in going from a 1-D to 2-D task. These findings are not surprising since the 2-D task's mean target velocities and displacements are greater than those required by the 1-D tasks. Thus, the increase in difficulty, and corresponding decrease in performance, may be due to these increased velocities and displacements and not the 2-D nature of the task. A 2-D task with the same horizontal and vertical target signals will have a mean velocity and displacement $\sqrt{2}$ times that of the corresponding 1-D task (if both signals are not in phase the scale factor will not be $\sqrt{2}$ but will still be greater than one) and hence it is likely that the MAEs observed in the 2-D task will be significantly greater than those of the corresponding 1-D tasks. For the target signal used in this study the 2-D task's average target displacement was 1.55 times that in the 1-D task and average velocity was 1.64 times that of the 1-D task. Assuming that is is possible to model the human S-M system as a linear system, which should be a reasonable first approximation here as the target signal should not significantly stress the S-M system, the absolute error observed on tracking tasks should increase in proportion to the task difficulty. As the task difficulty is a function of the average velocity and displacement it should be possible to divide the MAEs observed on the 2-D task by the above scale factors to assess the effect of the increased velocity and displacement of the 2-D task. On dividing $MAE2D_R$ by the factors it was found that for the normal subjects the results were close to the obtained $MAE1D_X$ and $MAE1D_Y$ ($MAE2D_R$ = 4.69 mm, 4.69/1.55 = 3.03 mm, 4.69/1.64 = 2.68 mm, cf. $MAE1D_X$ = 2.99 mm and $MAE1D_y$ = 3.09 mm) whereas for the Parkinsonian subjects the results were slightly above the observed $MAE1D_X$ and $MAE1D_Y$ ($MAE2D_R$ = 11.46 mm, 11.46/1.55 = 7.39 mm, 11.46/1.64 = 6.98 mm, cf. $MAE1D_X$ = 6.12 mm and $MAE1D_y$ = 6.39 mm). Thus, the observed deterioration in normal performance may be due to the increased physical difficulty of the 2-D task (i.e. higher velocity and displacement of target) and not increased dimensionality. Similarly, as noted above, the degradation in Parkinsonian performance when tracking in 2-D may be due to the increased difficulty simply in terms of motor execution. However, a reduction in the ability of the 'central-executive' (Baddeley *et al.*, 1986; Dalrymple Alford *et al.*, 1994) (central processing unit) of Parkinsonian subjects to appropriately allocate resources in tasks involving multiple degrees of freedom or interference between dimensions may also be, at least, contributing factors. To examine whether 2-D tracking performance is inferior to 1-D due to the increase in dimensionality or difficulty, another study comparing 2-D tracking performance to 1-D performance was performed (see §10.2). In this follow up study the 2-D target signals were scaled so that the resultant signal had the same mean velocity and displacement as the 1-D tasks' target signals. Full balancing of order effects also allowed elimination of the possibility that previous exposure to the 2-D task in this study allowed subjects to perform better on the 2-D task than would otherwise have been the case.

It is also interesting to note that on comparing the 1-D performance measures to those from the same coordinate of the 2-D task it was found that the normal tracking performance did not degrade on going to 2-D, while Parkinsonian performance did. These results are consistent with those of Chernikoff *et al.* (1960) and Todosiev *et al.* (1967), who found no degradation in the tracking performance of normal subjects, in the common coordinate, on going from 1-D to 2-D. It is important to note that as the author's 2-D tracking tasks are sufficiently sensitive that they effectively have no performance ceiling and are therefore able to detect even small differences between 1-D and 2-D performance. As explained in §8.3.4 the author has reservations about using the common coordinate technique to compare 1-D and 2-D tracking ability (due to the differences in the overall difficulty of the 1-D and 2-D tasks not associated with the increased dimensionality and the questionable belief that a 2-D task is equivalent to two 1-D tasks) and the above results confirm the suspect nature of this comparison. The results

from this comparison do, however, appear to add further weight to the assertion that the 2-D tracking tasks may provide new or more sensitive information on abnormal S-M function than 1-D tracking tasks. That is, on going from a 1-D to a 2-D tracking task this comparison has demonstrated that there is a significant degradation in Parkinsonian performance while no such degradation is observed for the normal subjects.

With one exception, there was no difference in performance on the 2-D tracking tasks, for random or step targets, between the vertical and horizontal directions. Superior performance was, however, observed on the vertical target entry times for the step task. As no difference was observed on all the other performance measures for the step task, the differences must be due to superior secondary corrective movements in the vertical direction. Whether this directional bias is due to a vertical visuospatial superiority, a fore-aft superiority on the joystick, or some other basis is unknown and is complicated by a lack of directional bias on other ballistic performance measures on the step task or on non-ballistic measures on the random task. Some evidence was seen for the supposition that subjects adopt a tracking strategy which is identical in both the vertical and horizontal directions during the primary ballistic movement but which diverges to an underdamped and overdamped strategy for the vertical and horizontal directions respectively when executing the secondary corrective movements to align with the target. The author is unaware of a physiological or psychomotor basis which could account for this conditional directional bias.

Improvement in the step task performance measures with prior knowledge of target position was expected for the normal subjects. Likewise, that Parkinsonian subjects demonstrate a similar ability is in agreement with previous studies that have shown Parkinsonian subjects can use prior information to develop an advance motor plan and thereby improve their performance, via a decrease in the observed lag, on smooth pursuit tracking tasks (Bloxham *et al.*, 1984; Day *et al.*, 1984; Jones and Donaldson, 1989). The improved performance in the current study was due, at least in part, to significantly shorter reaction times. This finding is contrary to other studies involving step stimuli in which only small improvements were seen on reaction times (Flowers, 1978; Bloxham *et al.*, 1984). The tasks used by Bloxham *et al.* (1984), Flowers (1978) and the author were, however, quite different (i.e. Bloxham's was a choice reaction timer and Flower's was a 1-D step tracking task) and it is therefore difficult to determine why the subjects in the previous studies showed only marginal improvements in reaction times. The only noticeable difference between the studies is that the accuracy of the entire step response was of interest here, whilst in the other studies this was either not measured or not considered. Thus in the previous studies the subjects may have concentrated solely on their reaction times whilst in this study the subjects had to also concentrate on the accuracy of the response. Overall, results from this study support the growing evidence that Parkinsonian subjects can use prior knowledge to form an advance motor plan to improve their performance. Furthermore, by artificially removing the effect of motor deficits through additive and proportional models of performance improvement, the Parkinsonian subjects were shown to be *unimpaired* in their ability to benefit from advance spatial information - at least for a step target.

Overall, the study has confirmed that the 2-D tasks can provide valid and valuable information on abnormal and normal S-M function. The results here also demonstrate that the 2-D tracking task can provide different information on the human S-M system in its abnormal state since there was a noticeable difference in the tracking performance of Parkinson's subjects on 1-D and 2-D tracking tasks.

10.2 STUDY TWO: AN EVALUATION OF NORMAL S-M FUNCTION IN 1-D AND 2-D

10.2.1 Objectives

The study described in §10.1, which investigated various aspects of 2-D tracking performance, was unable to conclusively determine whether the tracking performance of normal subjects degrades on going from a 1-D to a 2-D tracking task. The inconclusive result was due to the difficulty in excluding the possibility that the increased displacement and velocity of the 2-D task, in comparison to that of the 1-D task, was responsible for any degradation in tracking performance on going from 1-D to 2-D and the inability to exclude the possibility that the results from the study may have been confounded by the subjects' previous exposure to the 2-D tasks prior to the commencement of the study. This second study was therefore conducted to determine in a definitive manner whether normal subjects find 2-D tracking more difficult than 1-D tracking. It was also of interest to determine whether any difference between 1-D and 2-D tracking performance was exacerbated with increasing age or was different for males and females.

More specifically, the aims of the second study were to answer the following questions:

- When tracking a 2-D signal comprising two unscaled 1-D signals is there any difference in the overall measures of tracking performance between the 2-D and 1-D tasks?
- When tracking a 2-D signal which has been scaled to give it the same difficulty, in terms of average velocity and displacement, as the 1-D signals from which it was constructed, is there any difference in the overall measures of tracking performance between the 2-D and 1-D tasks?
- If a deterioration is seen in 2-D tracking performance is this deterioration due to the 2-D nature of the task or the increased average displacement and/or velocity of the 2-D task?
- Is tracking the unscaled 2-D signal more difficult than tracking the scaled 2-D signal and, if so, are any differences related to the ratio of the two tasks' velocities?
- Is tracking targets moving in a horizontal direction easier than tracking targets moving in a vertical direction ?
- Where significant differences are found with respect to the above questions, are the differences dependent on age or gender?

10.2.2 Experimental Design

10.2.2.1 Subjects

Sixteen subjects (eight male and eight female) who had no previous history of any significant neurological or musculoskeletal disorder participated in this study. The subjects had no previous exposure to tracking tasks of the type used in this study (i.e. tracking 2-D signals with a joystick). All subjects had a visual acuity of 6/9 or better and were right handed. Four of the subjects in each gender group were aged between 20-40 yrs (mean age 25 yrs) while the remaining four subjects were aged between 55-75 yrs (mean age 63 yrs). A demographic breakdown of the sixteen subjects is presented in Table 10.3

SUBJECT DATA

Subject	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Age (yrs)	62	58	66	62	63	72	56	63	23	24	23	31	25	33	22	21
Gender	F	F	F	F	M	M	M	M	F	F	F	F	M	M	M	M
Vision	6/5	6/4	6/6	6/6	6/4	6/6	6/5	6/4	6/4	6/4	6/5	6/4	6/5	6/5	6/5	6/5

Table 10.3 Gender, age, and visual acuity of subjects.

10.2.2.2 Experimental Setup

The apparatus and test used in this study were the 1-D and 2-D random pursuit tracking tasks described in §9.1.1. The experimental setup and configuration of these tasks was almost exactly the same as that described in §10.1.2.2 for the preceding study. The exceptions were that the tasks used in this study had a duration of 70 s and a scaled version of the 0.2 Hz bandwidth random 2-D target signal used in the previous study was produced for use in this study. The scaled 2-D random signal was produced by scaling the random 1-D signals, by an appropriate scale factor, so that the 2-D signal constructed from the 1-D signals and the original 1-D signals had approximately the same average displacement and velocity. The original 1-D target signals had an average displacement and velocity of 28.0 mm and 21.3 mm/s respectively and after scaling the resulting 2-D signal had an average displacement and velocity of 28.0 mm and 23.5 mm/s respectively. Thus, the difficulty of the scaled 2-D task and 1-D tasks, in terms of the average displacement and velocity, are almost equivalent. The scaled 2-D task was made marginally harder than the 1-D tasks, as opposed to slightly easier, so that if no significant difference was found between 1-D and 2-D performance, this could not be attributed to the scaled 2-D task being easier than the 1-D tasks.

10.2.2.3 Experimental Procedure

The group of sixteen experimental subjects were divided up into four subgroups each comprising two male subjects (1 20-40 yrs, 1 55-75yrs) and two female subjects (20-40 yrs, 55-75yrs). The subgroups were chosen by matching the youngest of the remaining subjects to the eldest, so as to minimise any differences between the mean ages of the subgroups. The tasks the subjects were required to complete in this study were the 2-D random task, scaled 2-D random task, 1-D vertical random task, and 1-D horizontal random task. As in the previous study each of the subgroups performed the tasks in a different order so that any order effects could be balanced out. That is, the order of the tasks for each subgroup was arranged so that half the subjects did a scaled 2-D task before a unscaled 2-D task, half the subjects did one 1-D task before a scaled 2-D task, half the subjects did one 1-D task before a unscaled 2-D task, and half the subjects did a 1-D vertical task before the 1-D horizontal task. The above ordering should have eliminated any order effects in the comparison of 1-D *vs* unscaled 2-D performance, 1-D *vs* scaled 2-D performance, scaled *vs* unscaled 2-D performance and 1-D vertical *vs* 1-D horizontal performance. The experimental design should also have minimised any bias in the results that could exist due to a difference in handedness, age range or gender. The design also allowed the subjects to be split, with regards to age and sex, into two further subgroups so that the effects of these factors could be investigated.

10.2.2.4 Analysis

The results from the study were analysed using Wilcoxon matched pairs comparisons for the reasons outlined in §10.1.2.4. All the statistical tests were two-tailed except for comparisons of the entire group's 1-D and 2-D tracking performance.

10.2.3 Results

Is there any difference in performance between tracking a 2-D target signal and the 1-D target signals from which the 2-D signal was constructed? On comparing the overall MAEs presented in Table 10.4 for the unscaled 2-D task and 1-D tasks it can be seen that tracking performance on the 2-D task was inferior to that observed for the 1-D task. This inferiority in tracking performance on the 2-D task, in comparison to that observed for the 1-D tasks, is also evident in the lags which are much greater for the 2-D task than the 1-D tasks. In the horizontal direction there was no difference between the gains for the two tasks, while in the vertical direction there was a relatively small, but significant, difference.

If a 2-D target signal has been scaled to give it the same difficulty (i.e. same average velocity and displacement) as the 1-D signals from which it was constructed, is is there any difference in performance between tracking the 2-D target and the 1-D targets? As can be seen in Table 10.4 the overall MAE obtained from the scaled 2-D task was larger than either of the MAEs obtained from the two 1-D tracking tasks. This indicates that tracking the scaled 2-D signal was more difficult than tracking either of the 1-D target signals. The lags were also greater for the 2-D task than the 1-D tasks confirming the inferior performance when tracking the scaled 2-D signal as opposed to the 1-D signals. The gains for the two sets of tasks were also significantly different, although the differences in the two directions were of opposite sign and relatively small.

COMPARISON OF 2-D AND 1-D TRACKING PERFORMANCE

Performance Measure	2-D Mean	1-D Mean	Δ %	p
Unscaled 2-D Task				
$MAE2D_R$ *vs* $MAE1D_X$	7.0 mm	3.6 mm	94%	<0.001
$MAE2D_R$ *vs* $MAE1D_Y$	7.0 mm	3.5 mm	100%	<0.001
$Lag2D_X$ *vs* $Lag1D_X$	149 ms	106 ms	41%	<0.01
$Lag2D_Y$ *vs* $Lag1D_Y$	265 ms	108 ms	145%	<0.01
$Gain2D_X$ *vs* $Gain1D_X$	135	136	-1%	NS
$Gain2D_Y$ *vs* $Gain1D_Y$	120	118	2%	<0.01
Scaled 2-D Task				
$MAE2D_R$ *vs* $MAE1D_X$	5.0 mm	3.6 mm	39%	<0.001
$MAE2D_R$ *vs* $MAE1D_Y$	5.0 mm	3.5 mm	43%	<0.001
$Lag2D_X$ *vs* $Lag1D_X$	151 ms	106 ms	42%	<0.01
$Lag2D_Y$ *vs* $Lag1D_Y$	180 ms	108 ms	67%	<0.1
$Gain2D_X$ *vs* $Gain1D_X$	142	136	4%	<0.05
$Gain2D_Y$ *vs* $Gain1D_Y$	117	118	-1%	<0.05

Table 10.4

With either of the 2-D target signals is the tracking performance observed in the coordinates common to both the 1-D tasks different to that observed for the 1-D tasks?

For the unscaled 2-D task the errors in each coordinate were greater than those for the 1-D task ($MAE2D_X$ *vs* $MAE1D_X$: 4.1 mm *vs* 3.6 mm, 14%, $p<0.01$; $MAE2D_Y$ *vs* $MAE1D_Y$: 4.7 mm *vs* 3.5 mm, 34%, $p<0.01$) whereas for the 2-D scaled task the errors in each coordinate were less than those for the 1-D task ($MAE2D_X$ *vs* $MAE1D_X$: 3.1 mm *vs* 3.6 mm, -15%, $p<0.01$; $MAE2D_Y$ *vs* $MAE1D_Y$: 3.2 mm *vs* 3.5 mm, -9%, $p<0.05$)

Does the extra dimensionality of the 2-D task contribute to its increased difficulty over the 1-D task? The deterioration in tracking performance on going from a 1-D to a 2-D tracking task, with the same average displacement and velocity as the 1-D task, occurred because of the extra dimensionality of the 2-D task. This was verified by the increases in the overall error measures and lags on going from the 1-D tasks to the 2-D scaled task as shown in Table 10.4. The increase in the lags on going to 2-D occurred because of increased delays within the S-M system or a reduction in the ability of the system to reduce these delays by using prediction. The increase in lags cannot be due to any limitation in motor function as the 1-D and 2-D targets have the same average velocity and displacement. Going from 1-D to 2-D required an average 59 ms (((151-106)+(180-107))/2) of extra processing time and contributed on average an extra 1.5 mm (((5.0-3.6)+(5.0-3.5))/2) to the overall error.

Does the deterioration seen in 2-D tracking performance occur equally in each coordinate? The deterioration in performance on going from a 1-D task to either the scaled or unscaled 2-D tasks occurred equally in each coordinate as the additive and proportional models of performance degradation, comparing ($MAE2D_X - MAE1D_X$) to ($MAE2D_Y - MAE1D_Y$), show for the unscaled 2-D task (Additive: 0.6 mm *vs* 1.2 mm, NS; Proportional: 10% *vs* 14%, NS) and scaled 2-D task (Additive: -0.5 mm *vs* -0.3 mm, NS; Proportional: -17% *vs* -22%, NS). A similar result is also seen for additive and proportional models of performance degradation in terms of the lags, as no difference is observed for either the unscaled 2-D task (Additive: -44 ms *vs* -160 mm, NS; Proportional: -27% *vs* -28%, NS) or scaled 2-D task (Additive: -46 ms *vs* -75 ms, NS; Proportional: -31% *vs* -9%, NS).

Is tracking the unscaled 2-D signal more difficult, at least in terms of absolute errors, than tracking the scaled 2-D signal? Tracking the unscaled 2-D signal was more difficult than tracking the scaled 2-D signal. That is, the MAEs from the unscaled 2-D task are much larger than those obtained from the scaled 2-D task (Resultant: 7.0 mm *vs* 5.0 mm, 29%, $p<0.001$; X: 4.1 mm *vs* 3.1 mm, 25%, $p<0.001$; Y: 4.7 mm *vs* 3.2 mm, 32%, $p<0.01$). There is a significant difference between the lags obtained from the two tasks in the vertical direction (265 ms *vs* 180 ms, 32%, $p<0.05$) but not in the horizontal direction (149 ms *vs* 151 ms, 0%, NS). The large difference between the mean lags in the vertical direction appears due to two older subjects who had considerably larger lags in this direction than the rest of the subjects. Paradoxically, however, the majority of the subjects' ($\frac{12}{16}$) unscaled task vertical lags were *less* than those from the scaled task and so the significant p-value above is, in fact, due to the lags from the unscaled task being, on average, less than those from the scaled task (i.e. not the other around as the means suggest). In fact the median lags in the vertical direction for the unscaled and scaled tasks are 99 ms and 133 ms respectively which are almost equivalent since the cross-correlation techniques used to compute these values has a resolution of 34 ms (the horizontal direction medians were 132 ms and 132 ms). There is also no conceivable reason why the lags from the scaled task should be greater than those from the unscaled task (i.e. any additional processing required should be in the direction of the more difficult unscaled task) and as the difference between the horizontal unscaled and scaled lags are small (i.e. $\leq$ 68 ms) for eight out of the sixteen subjects it seems reasonable to assume that the result seen in the vertical direction is a statistical anomaly. It is also worth noting, as is reported latter, when a superiority is seen in one direction it is usually found that the vertical performance is superior

to the horizontal performance and not the other way around as, at first appearances, appears to be the case here for the unscaled task. Thus, on considering the overall result seen in the horizontal direction and the discussion above regarding the result in the vertical direction, the author believes it is reasonable to conclude that the amount of processing required to complete the scaled and unscaled tasks is equivalent (i.e. in each direction the lags from the unscaled and scaled tasks should be approximately equal).

The difference between the gains, while of the same sign for most of the matched pairs, was negligible (X: 136 *vs* 136, 0%, $p<0.05$: Y: 120 *vs* 118, 0% , $p<0.05$). In the horizontal direction the median gains for the unscaled and scaled tasks were 133 and 137 respectively, while in the vertical direction they were 114 and 111 respectively. In the vertical direction the majority of the subjects ($\frac{11}{16}$) had gains from the unscaled task that were greater or equal to those from the scaled task and in the horizontal direction the majority ($\frac{12}{16}$) had gains for the scaled task that were greater or equal to those in the unscaled task.

Are the differences observed in tracking the unscaled 2-D signal and scaled 2-D signal related to the ratio of the average velocities of the two tasks? The average target velocities in the unscaled and scaled 2-D tasks was 35.18 mm/s and 23.53 mm/s respectively, giving a ratio of 1.50. To determine whether the differences in performance on the two tasks were directly related to this average velocity ratio the MAEs obtained from the unscaled 2-D task were divided by 1/1.50 and compared to those obtained from the scaled 2-D task. On comparing the scaled MAEs from the unscaled 2-D task to those from the scaled 2-D task it was found that, while similar, there still tends to be a difference between the two sets of MAEs ($MAE2D'_R$ *vs* $MAE2D_R$: 4.7 mm *vs* 5.0 mm, -6%, $p<0.1$; $MAE2D'_X$ *vs* $MAE2D_X$: 2.7 mm *vs* 3.1 mm, -13%, $p<0.05$; $MAE2D'_Y$ *vs* $MAE2D_Y$: 3.2 mm *vs* 3.2 mm, -2%, NS).

Did the informal practice before the beginning of the previous study bias its findings? Clearly the informal practice before the beginning of the previous study did bias the results comparing 1-D and 2-D tracking ability. The subjects from the previous study had exposure to an unscaled 2-D tracking task before participating in the previous study. The results showed that there was no significant difference between normal 1-D and 2-D tracking ability when comparing the tracking error in each coordinate and when comparing the 1-D error to the 2-D error scaled by the ratio of the mean task velocities. In this study, the unscaled 2-D task performance in each coordinate was consistently greater than that in the corresponding coordinate in the 1-D task ($MAE2D_X$ *vs* $MAE1D_X$: 4.1 mm *vs* 3.6 mm, 14%, $p<0.01$; $MAE2D_Y$ *vs* $MAE1D_Y$: 4.7 mm *vs* 3.5 mm, 26%, $p<0.01$) and the 2-D resultant error scaled by the velocity ratio was also greater than the 1-D errors ($MAE2D'_R$ *vs* $MAE1D_X$: 4.7 mm *vs* 3.6 mm, 23%, $p<0.05$; $MAE2D'_R$ *vs* $MAE1D_Y$: 4.7 mm *vs* 3.5 mm, 26%, $p<0.001$). Hence, practice on the unscaled 2-D task immediately before the previous study enabled the subjects to perform better on the 2-D task in that study than would have otherwise been the case.

Is tracking targets moving in a horizontal direction easier than tracking targets moving in a vertical direction? On comparing performance in the horizontal and vertical directions of all the tasks it was found that tracking a target moving in the horizontal direction was no more difficult than tracking a target moving in the vertical direction. That is, in the horizontal and vertical directions respectively, there were similar MAEs and lags in the 1-D tasks (MAE: 3.6 mm *vs* 3.5 mm, 0%, NS; lag: 106 ms *vs* 107 ms, 0%, NS), unscaled 2-D task (MAE: 4.1 mm *vs* 4.7 mm, 14%, NS; lag: 149 ms *vs* 265 ms, 78%, NS) and scaled 2-D task (MAE: 3.1 mm *vs* 3.2 mm, 4%, NS; lag: 151 ms *vs* 180 ms, 19%, NS). The tracking strategy in the horizontal and vertical coordinates must, however, be different as the gains in

each coordinate are significantly different for all the tasks (1-D: 136 *vs* 118, 13%, $p<0.001$; 2-D Scaled: 136 *vs* 121, 11%, $p<0.001$; 2-D Unscaled: 136 *vs* 117, 13%, $p<0.01$).

Is there any difference in the tracking performance of the younger and older groups? The performance of the younger age group was superior to that of the older age group on all tasks. The superiority in the younger group's tracking performance can be seen in the smaller MAEs and lags on the unscaled 2-D task ($MAE2D_R$: 4.3 mm *vs* 9.7 mm, 55%, $p<0.01$; $Lag2D_X$: 99 ms *vs* 199 ms, 50%, $p<0.05$; $Lag2D_Y$: 62 ms *vs* 472 ms, 86%, $p<0.01$), scaled 2-D task ($MAE2D_R$: 3.6 mm *vs* 6.3 mm, 42%, $p<0.01$; $Lag2D_X$: 116 ms *vs* 186 ms, 38%, $p<0.05$; $Lag2D_Y$: 83 ms *vs* 282 ms, 70%, $p<0.05$) and 1-D tasks ($MAE1D_X$: 2.8 mm *vs* 4.3 mm, 33%, $p<0.01$; $MAE1D_Y$: 2.4 mm *vs* 4.6 mm, 49%, $p<0.01$; $Lag1D_X$: 70 ms *vs* 141 ms, 50%, $p<0.1$; $Lag1D_Y$: 58 ms *vs* 162 ms, 64%, $p<0.05$).

On comparing the older and younger age groups' degradation in performance on going from the 1-D tasks to the scaled 2-D task, it was found that this increased marginally with age. As can be seen in Table 10.5 the additive and proportional models of performance degradation, for the scaled 2-D task, show an increase in the MAE difference (i.e. $MAE2D_R - MAE1D_X$) when the 1-D horizontal task is used for comparison purposes. Conversely, when the 1-D vertical task is used for comparison purposes (i.e. $MAE2D_R - MAE1D_Y$) no significant increase in the performance degradation is observed with age. Table 10.5 also shows that the increase in performance degradation with age, on going from a 1-D to a 2-D task, is not observed in the lag and gains. This may indicate that the increased 1-D to 2-D differential with age may not be related to slower processing in the older group.

Comparison of the difference in performance on the unscaled 2-D task and 1-D task with regards to age (Table 10.5) shows that the difference between the 1-D and 2-D errors, when both the 1-D horizontal and vertical tasks are used for comparison purposes, increases with age when using the additive model. The proportional model does not show the same trend. As the previous comparison with the scaled 2-D task has shown only a marginal difference between the two age groups, with regards to the difference in the errors for the 1-D and 2-D tasks, the difference here must primarily be due to the increased velocity of the unscaled 2-D task. It also appears that the change in tracking strategy that must occur on going from a 1-D to a 2-D task does not differ markedly with age since no increase in the difference in the lags and gains was observed with age.

To determine whether performance in the horizontal coordinate of a task in comparison to that in the vertical coordinate changes with age the performance in each coordinate of each of the tasks has been compared for both age groups. For the older subjects there was no difference between the performance in the vertical and horizontal directions of the scaled 2-D task ($MAE2D_X$ *vs* $MAE2D_Y$: 3.7 mm *vs* 4.4 mm, 17%, NS), unscaled 2-D task ($MAE2D_X$ *vs* $MAE2D_Y$: 5.2 mm *vs* 7.0 mm, 26%, NS), and 1-D tasks ($MAE1D_X$ *vs* $MAE1D_Y$: 4.3 mm *vs* 4.6 mm, 17%, NS) although the tendency was for superior performance in the horizontal coordinate. Conversely, the performance of the younger subjects was superior in the vertical coordinate of the 1-D ($MAE1D_X$ *vs* $MAE1D_Y$: 2.8 mm *vs* 2.4 mm, 20%, $p<0.01$) and unscaled 2-D ($MAE2D_X$ *vs* $MAE2D_Y$: 3.1 mm *vs* 2.5 mm, 24%, $p<0.01$) tasks, but not the scaled 2-D task ($MAE2D_X$ *vs* $MAE2D_Y$: 2.6 mm *vs* 2.1 mm, 23%, NS). Differences in the lags were only observed in the results from the 2-D unscaled task for both the older (Horizontal *vs* Vertical: 198 ms *vs* 472 ms, -58%, $p<0.1$) and younger (Horizontal *vs* Vertical: 99 ms *vs* 62 ms, 60%, $p< 0.05$) age groups.

The difference between the two age groups with respect to horizontal and vertical performance is more convincingly demonstrated by looking at the difference in performance between

COMPARISON OF DIFFERENCES IN 2-D AND 1-D PERFORMANCE WITH AGE

Performance Measure	20-40 yrs	55-75 yrs	p
Scaled 2-D task: Additive Model			
$MAE2D_R - MAE1D_X$	0.8 mm	2.1 mm	<0.05
$MAE2D_R - MAE1D_Y$	1.3 mm	1.7 mm	NS
$Lag2D_X - Lag1D_X$	46 ms	46 ms	NS
$Lag2D_Y - Lag1D_Y$	-25 ms	120 ms	<0.1
$Gain2D_X - Gain1D_X$	7	6	NS
$Gain2D_Y - Gain1D_Y$	4	-5	NS
Scaled 2-D task: Proportional Model			
$MAE2D_R - MAE1D_X$	17%	27%	<0.1
$MAE2D_R - MAE1D_Y$	32%	24%	NS
$Lag2D_X - Lag1D_X$	35%	26%	NS
$Lag2D_Y - Lag1D_Y$	-9%	20%	NS
$Gain2D_X - Gain1D_X$	4%	4%	NS
$Gain2D_Y - Gain1D_Y$	2%	-4%	NS
Unscaled 2-D task: Additive Model			
$MAE2D_R - MAE1D_X$	1.5 mm	5.4 mm	<0.01
$MAE2D_R - MAE1D_Y$	1.9 mm	5.0 mm	<0.05
$Lag2D_X - Lag1D_X$	29 ms	58 ms	NS
$Lag2D_Y - Lag1D_Y$	4 ms	310 ms	<0.1
$Gain2D_X - Gain1D_X$	0	0	NS
$Gain2D_Y - Gain1D_Y$	2	5	NS
Unscaled 2-D task: Proportional Model			
$MAE2D_R - MAE1D_X$	33%	50%	<0.1
$MAE2D_R - MAE1D_Y$	46%	48%	NS
$Lag2D_X - Lag1D_X$	27%	27%	NS
$Lag2D_Y - Lag1D_Y$	0%	50%	<0.1
$Gain2D_X - Gain1D_X$	0%	0%	NS
$Gain2D_Y - Gain1D_Y$	1%	2%	NS

Table 10.5

the two coordinates by using the additive and proportional models of performance degradation. These show the difference between the younger and older groups' difference in horizontal and vertical performance on the scaled 2-D (Additive: 0.5 mm *vs* -0.7 mm, $p<0.01$; Proportional: 23% *vs* -12%, $p<0.01$), 2-D unscaled (Additive: 0.6 mm *vs* -1.8 mm, $p<0.01$; Proportional: 20% *vs* -25%, $p<0.05$), and 1-D (Additive: 0.5 mm *vs* -0.4 mm, $p<0.05$; Proportional: 18% *vs* -7%, $p<0.05$) tasks. These results confirm that the relationship between vertical and horizontal performance for the older age group is completely different to that for the younger age group.

Is there a difference between male and female performance? There was no significant differences between the performance of the male and female groups on the scaled 2-D ($MAE2D_R$: 4.4 mm *vs* 5.6 mm, 28%, NS), unscaled 2-D ($MAE2D_R$: 6.8 mm *vs* 7.2 mm, 6%, NS), and 1-D tasks ($MAE1D_X$: 3.5 mm *vs* 3.6 mm, 2%, NS; $MAE1D_Y$: 3.4 mm *vs* 3.6 mm, 8%, NS). Similarly, no difference between male and female performances were observed on comparing lags from the scaled 2-D task ($Lag2D_X$: 149 ms *vs* 149 ms, 0%, NS; $Lag2D_Y$: 249 ms *vs* 286 ms, 15%, NS), unscaled 2-D task ($Lag2D_X$: 145 ms *vs* 157 ms, 9%, NS; $Lag2D_Y$: 153 ms *vs* 211 ms, 38%, NS), and 1-D tasks ($Lag1D_X$: 116 ms *vs* 95 ms,

-18%, NS; $Lag1D_Y$: 99 ms *vs* 120 ms, 21%, NS).

Furthermore, no difference was found between male and female groups in the degradation in performance on going from a 1-D to a 2-D task. The data in Table 10.6 indicates, however, that there may be a tendency for greater performance degradation with the female subjects since the difference in the MAEs and lags are greater for the female subjects than the males.

COMPARISON OF THE DIFFERENCE IN 2-D AND 1-D PERFORMANCE WITH GENDER

Performance Measure	Male	Female	p
Scaled 2-D task: Additive Model			
$MAE2D_R - MAE1D_X$	0.8 mm	2.0 mm	NS
$MAE2D_R - MAE1D_Y$	1.0 mm	2.0 mm	NS
$Lag2D_X - Lag1D_X$	29 ms	62 ms	NS
$Lag2D_Y - Lag1D_Y$	54 ms	91 ms	NS
Scaled 2-D task: Proportional Model			
$MAE2D_R - MAE1D_X$	18%	26%	NS
$MAE2D_R - MAE1D_Y$	26%	30%	NS
$Lag2D_X - Lag1D_X$	27%	35%	NS
$Lag2D_Y - Lag1D_Y$	2%	9%	NS
Unscaled 2-D task: Additive Model			
$MAE2D_R - MAE1D_X$	3.3 mm	3.6 mm	NS
$MAE2D_R - MAE1D_Y$	3.4 mm	3.5 mm	NS
$Lag2D_X - Lag1D_X$	33 ms	54 ms	NS
$Lag2D_Y - Lag1D_Y$	149 ms	165 ms	NS
Unscaled 2-D task: Proportional Model			
$MAE2D_R - MAE1D_X$	41%	43%	<0.1
$MAE2D_R - MAE1D_Y$	47%	46%	NS
$Lag2D_X - Lag1D_X$	22%	33%	NS
$Lag2D_Y - Lag1D_Y$	22%	27%	NS

Table 10.6

The male and female groups were able to track equally well in both the horizontal and vertical directions with no difference in performance being found on the scaled 2-D task(Male: 2.8 mm *vs* 2.7 mm, 4%, NS; Female: 3.4 mm *vs* 3.7 mm, 10%, NS), unscaled 2-D task (Male: 4.1 mm *vs* 4.5 mm, 10%, NS; Female: 4.2 mm *vs* 4.9 mm, 16%, NS), and 1-D tasks (Male: 3.5 mm *vs* 3.4 mm, 5%, NS; Female: 3.6 mm *vs* 3.6 mm, 1%, NS).

10.2.4 Discussion

This second study was primarily conducted to determine in a definitive manner whether 2-D tracking is more difficult than 1-D tracking and, if so, whether the increased difficulty can be attributed to the multidimensional nature of the target. Results from this study have conclusively demonstrated that tracking in 2-D is appreciably more difficult than tracking in 1-D and that the multidimensional nature of the 2-D target is a major basis for this increased difficulty. The result demonstrating this drop in tracking performance came from the 2-D tracking task which had had its target signal scaled to give it the same average displacement and velocity as the 1-D tasks from which it was derived. The resultant MAEs from the 2-D task were on average 42 % larger than the MAEs from either the horizontal or vertical 1-D task indicating that

the drop in performance must have been due to the multidimensional nature of the 2-D task. When the 2-D task does not have the same average displacement and velocity as the 1-D tasks from which it was derived, the degradation in tracking performance comprises a component related to the increase in velocity in addition to the component associated with the increase in dimensionality. The evidence for this belief is the larger MAEs observed from the unscaled 2-D task in comparison to those observed from the scaled 2-D task.

Further evidence that the drop in performance on going to 2-D was because of the multidimensional nature of the task was the 55 % increase in the subjects' tracking lag. The increase in lag was similar in both coordinates. To determine how much of the increase in the MAE could be accounted for by the increase in the lag the author produced a series of simulated uncontaminated response signals and computed the resulting MAE as the lag was increased (i.e. response signals were generated using Eqn.(9.16) where $c[n] = 0$). As indicated in §10.2.3, going from 1-D to 2-D leads to an extra 59 ms time delay and an extra 1.5 mm overall error. The simulation suggested that a lag of 59 ms would contribute approximately 1.1 mm to the overall error (c.f. the observed 1.5 mm), suggesting that the increased time delay is the largest contributor to the increase in the MAE (i.e. $\frac{1.1}{1.5} \Rightarrow 73\%$). This increased time delay could be due to a reduction in the ability of the S-M system to reduce the inherent delays within the system (i.e. physical propagation delays along nerves and information processing delays) by using prediction to develop an anticipatory response or an increase in the information processing delays because of an increase in the amount of information that must be processed. The remaining 27 % of the increase in the MAE, which has yet to be accounted for, could be due to factors such as the subjects performing open-loop movements to 'catch-up' to the target when they perceive that their tracking error is becoming large, a reduction in the accuracy with which the S-M system can predict the target's future position, and a reduction in the accuracy of the movement plans. Further confirmation of the reduction in the ability of the S-M system to use prediction to offset the inherent time delay or of the increased information processing demands on going to 2-D was obtained from comparison of lags on the scaled 2-D and unscaled 2-D tasks. This comparison showed no appreciable difference between the lags from the two tasks, despite the greater physical difficulty of the unscaled 2-D task. Although this comparison of lags on the two 2-D tasks was overshadowed by the result indicating a significant difference between the lags in the vertical direction (see §10.2.3), the difference is in a direction opposite to that which could disprove the above belief (i.e. the lags on the scaled 2-D task tend to be larger than those on the unscaled 2-D task in the vertical direction).

In terms of the model of the S-M system described in §8.1 it appears that the increased time delay on going to 2-D could be due to a reduction in the performance of the sensory-analysis and the response planning/selection subsystems. The sensory-analysis subsystem is responsible for target prediction (see (Neilson *et al.*, 1992)) and it is possible that the increased delay is due to the reduced ability of this system to provide accurate predictions, far enough into the future, to allow the inherent delays within the S-M system to be reduced to the same extent that they were reduced in the 1-D case. The portion of the sensory analysis subsystem that processes the information provided by the sense organs is unlikely to be effected by the multidimensional nature of the target as it is already processing multidimensional information from the sense organs and does so in a parallel manner. If the 1-D tasks require the full resources of the response planning/selection subsystem, it is possible that increased information processing delays in this system, on going to 2-D, might also be contributing to the increased time delay. It might be necessary for this system to allocate its resources in a sequential manner to allow all the additional information required to complete the task to be processed, hence the time delay. The response planning/selection subsystem is also the system that decides on the most

appropriate type of response and it is therefore the system that is responsible for any open-loop movements which are executed to catch-up to the target and for any reduction in the accuracy of the movement plans. The performance of the response execution system should not degrade on going to 2-D as it must always carry out its processing in a 3-D frame of reference independent of the dimensionality of the stimulus. That is, the human functions in a 3-D environment, with any movement requiring a plan to move from one point in 3-D space to another point in 3-D space; as a consequence, the delay in this subsystem should not be effected by the dimensionality of the stimulus.

As stated above, the degradation in performance that occurs on going from the 1-D tasks to the unscaled 2-D task is predominantly composed of two components: one due to the increase in average target velocity and the other due to the time delay increase and decrease in tracking accuracy that occurs because of the increase in task dimensionality. In an effort to describe mathematically how performance degrades on going from a 1-D to a unscaled 2-D task, the author developed the preliminary expression

$$MAE2D'_R(MAE1D) = V_{SF}[MAE1D + \epsilon_{12D}], \tag{10.1}$$

which describes the 2-D task's MAE in terms of the MAE observed on the corresponding 1-D task, the error ϵ_{12D} describing the degradation associated with the increased time delay and increased tracking error in 2-D, and the velocity scale factor V_{SF} which is the ratio of the average velocities of the 2-D and 1-D target signals. For the current study ϵ_{12D} was 1.5 mm and V_{SF} was 1.5 (see §10.2.3). On substituting these values and the obtained 1-D MAEs into Eqn.(10.1) it was found that the expression could predict with a reasonable accuracy the MAE obtained on the unscaled 2-D task ($MAE2D_R$ *vs* $MAE2D'_R(MAE1D_X)$: 7.0 mm *vs* 7.7 mm, 10 %, $p < 0.1$; $MAE2D_R$ *vs* $MAE2D'_R(MAE1D_Y)$: 7.0 mm *vs* 7.5 mm, 7 %, $p < 0.1$). It needs to be noted that Eqn.(10.1) is a simplistic first attempt at describing the $MAE2D_R$ in terms of the observed $MAE1D$ and an expression would have to be derived for ϵ_{12D} in terms of its contributing components for Eqn.(10.1) to be useful. The time delay which is responsible for the majority of ϵ_{12D} might be a constant independent of the target signal (this requires further investigation) and, if this is the case, the major component comprising ϵ_{12D} could then simply be calculated from the 1-D target signal and a delayed version. It might also be possible to develop models describing how any open-loop catch-up movements, target prediction inaccuracies, and response planning/selection inaccuracies contribute to the remaining major component of ϵ_{12D}. The results in §10.2.3, concerning the scaling of the unscaled 2-D results to obtain the scaled 2-D results, also show that the relationship between the MAE and a task's average velocity is not necessarily a simple linear relationship devoid of any other contributing factors (e.g. the effects of ageing). Thus, while Eqn. (10.1) seems to be a reasonably good first approximation that is backed up by some strong experimental findings it needs further refining to be useful.

On comparing the errors in each coordinate of the unscaled 2-D tracking task to the errors in the corresponding coordinate of the 1-D tasks it was found that the errors from the 2-D task were greater than those from the 1-D task. This should be compared to the result from the previous study which found no difference in the performance of the normal group on the two tasks. The reason no difference could be found in the previous study is most likely due to the subjects' previous exposure to the 2-D task. It was also found on conducting this comparison with the MAEs from the scaled 2-D task and the 1-D tasks that the errors from the 1-D tasks were greater than those from the 2-D task. The set of results obtained in the current study for the unscaled 2-D task agree with those results of Garvey and Taylor (1959) and Ziegler (1968) presented in §8.3.4, while the results from the scaled 2-D task agree with those of Chernikoff

et al. (1960) and Todosiev *et al.* (1967). From their results Garvey and Taylor (1959) and Ziegler (1968) concluded that the two discrete tasks were competing for the same limited S-M resources and performance in the common coordinate degrades because the finite capacity can no longer be dedicated to one task. As explained in §8.3.4 all of these authors believed that a 2-D task could be considered as two independent tasks being performed at the same time. The variability in the above results, which indicate that the average target velocity of the 1-D and 2-D tasks does influence the result, adds further weight to the author's belief that undertaking comparisons of 1-D and 2-D tracking performance in common coordinates is a questionable technique for evaluating how the S-M system allocates resources.

On separating the results from the 20-40 yrs age group from those of the 55-75 yrs age group it was found, for all the tracking tasks, that the average errors and lags of the older group were consistently greater than those of the younger age group. The increase in lags demonstrates that the information processing capability, in terms of speed or capacity, of the older subjects is reduced and/or the speed with which they can initiate and execute movements has been reduced. The overall increase in tracking error with age is probably directly related to this observed increase in the lags, although there may also be a contribution from other components of the S-M system that are known to deteriorate with age such as visual acuity, visual perception, etc. (Jones *et al.*, 1986). Comparison of the difference between the 1-D and 2-D tracking performance of the older and younger groups demonstrated that the ageing process did not affect the extra processing required by the 2-D task. The evidence for this was that for both the scaled and unscaled 2-D tasks there was no real consistent difference between the two groups differences in the 2-D and 1-D lags, indicating that the time required to perform the 2-D processing did not increase significantly with age. There was, however, a more consistent increase in the difference between the 1-D and 2-D tracking errors, particularly for the unscaled 2-D task, indicating that the increased velocity of the unscaled task may be the factor responsible for the degradation of the older subject's performance. These results agree with those of Jones *et al.* (1986) who found on 1-D random and step tasks that tracking performance deteriorated with age. Jones *et al* concluded, via tests of component S-M function, that the deterioration is most likely due to an overall reduction in motor speed as opposed to a reduction in central processing performance. Other authors (Erikesen *et al.*, 1970; Botwinick, 1981) have found, however, that older subjects require more time to perceive and interpret visual stimuli. This study has shown that the older subjects do not require significantly more time than the young group to comprehend 1-D as opposed to 2-D target movement.

On comparing the performance of the male and female subjects it was found that there was no difference between the two groups performance on any of the tasks. Similarly, on going from 1-D to 2-D the degradation in performance for the females was no different to that for the males, indicating that neither groups spatial comprehension is better than the others. These results are in disagreement with previous studies by Jones *et al* (Jones and Donaldson, 1981; Jones *et al.*, 1986) who found that male performance on pursuit tracking tasks, with a steering wheel as the motor output device, was superior to that of female performance. As previous studies confirmed that males are stronger than females on grip strength (Potvin *et al.*, 1973; Agnew and Maas, 1982; Fraser and Benton, 1983; Jones *et al.*, 1986) and arm strength tasks (Potvin *et al.*, 1973; Jones *et al.*, 1986) than females, the results showing that there is no difference between male and female performance on any of the tasks implies that superior strength cannot be used to advantage when tracking smooth band-limited targets (at least for the low bandwidth (0.21 Hz) target signals used here). Similarly, there appears to be no major sociocultral differences that could give either group an advantage on joystick tracking tasks. Conversely, this may not be true for steering wheel tracking, particularly in older age groups,

where males may have more driving experience than females.

For the overall group there was no difference between vertical and horizontal performance on the random pursuit tracking tasks. Similarly, no superiority in performance in either direction was observed when the overall group was split with regards to gender. A difference in performance was, however, found between the young and old group. The young groups' vertical performance was superior to its horizontal performance whereas no such superiority was observed for the older age group, although there was a tendency for superiority in the horizontal direction. This is consistent with the previous study §10.1, which consisted solely of an older age group, where no difference was observed between smooth tracking performance in the two coordinates. The younger groups' superiority may or may not be related to the, as yet unexplained, superiority of the secondary corrections in the vertical direction of the step task noted for older subjects in the previous study. The current results may reflect some superiority in the strength or controllability at high speed of the muscle groups required to produce fore-aft movements over those muscles used to produce left-right movements. It is unclear whether this directional superiority in motor control diminishes with age or is simply due to a change in tracking strategy (that may occur because of the general slowing down of the S-M system with age) resulting in no benefit being made of any superior control in the vertical/fore-aft direction. For example, older subjects may slow all there movements down to that which can be achieved by the muscle group with the poorest performance.

In summary, the results demonstrate that tracking performance does degrade on going from a 1-D to a 2-D tracking task. When the 2-D task has the same average displacement and velocity as the 1-D task, the majority of the degradation is due to a reduction in the ability of the S-M system to offset the inherent time delay in the system or increased delays because of increased information processing demands. If the 2-D task has a larger average displacement or velocity these factors will also contribute to the degradation in performance. These results indicate that the use of 2-D tasks, in conjunction with 1-D tasks, would be useful for investigating multidimensional S-M functioning and how this is disrupted in certain neurological disorders. These results also demonstrate that 2-D tasks can provide new and more sensitive measures of S-M function than 1-D tasks and there are therefore benefits to be gained from using 2-D tracking tasks in addition to 1-D tasks. It is also worth noting that it is still uncertain at this stage whether the additional degradation seen in Parkinsonian performance on going from 1-D to 2-D in the previous study was due to a deficit of a multidimensional nature or due to the increased average velocity of the target. With regards to age and gender, there does not appear to be any major difference in the way in which tracking performance degrades on going from a 1-D to a 2-D task. It does appear, however, that subjects in the 20-40 yrs age range have superior performance in the vertical coordinate of the 2-D task while no superiority is observed for older subjects (55-75yrs).

CHAPTER 11

CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH - PART II

In Part Two of this thesis two 2-D tracking tasks developed to allow sensory-motor (S-M) function to be assessed in a multidimensional environment are described. Also described and discussed are the results from studies conducted to investigate various aspects of 2-D tracking and tracking tasks in general. Specifically, the results from two studies, conducted to verify the expectation that 2-D tracking tasks can provide different and additional information on normal or abnormal S-M function to that provided by 1-D tasks, are presented. In addition, results are presented from a study conducted to determine why the technique originally used to estimate the time delay between the tracking task target signal and the S-M response signal, provided inaccurate estimates. The latter study was extended to include an investigation of the accuracy of a number of time delay estimation techniques that can be used in conjunction with tracking tasks.

In Chapter 8, the S-M system and the function of its major components are described in terms of a simple model. Also described are the various types of tests which have been used to evaluate the performance of the S-M subsystems or the performance of the system as a whole. This discussion indicates that tracking tasks are an extremely valuable means of assessing S-M function as they objectively provide quantitative measures of performance and information on S-M system strategies. It is also indicated that there maybe benefits to be gained from using 2-D tracking tasks in preference to 1-D tasks. The two 2-D tracking tasks developed by the author are then described in Chapter 9. Also presented in Chapter 9 are the results from the study used to evaluate the accuracy of the time delay estimation techniques that can be used with tracking tasks. In Chapter 10, the two studies comparing 1-D and 2-D tracking tasks are described and the results demonstrating that there are benefits to be gained from using 2-D tasks are presented. The original work presented in Part Two is contained within Chapters 9 and 10.

The aim of this chapter is to present the conclusions that have been drawn from the results, and subsequent discussions, in Chapters 9 and 10 and to describe those areas of research which are worthy of further work. The conclusions are presented in §11.1 and the suggestions for further work are given in §11.2.

11.1 CONCLUSIONS

Results from the studies described in Chapter 10 have shown that the two 2-D tracking tasks, described in §9.1, are a valuable addition to the upgraded S-M test battery at Christchurch Hospital. Specifically, the results presented in §10.1 indicated that the 2-D tracking tasks can

demonstrate and quantify the deficits known to occur in the Parkinsonian S-M system. The studies have also shown that the 2-D tasks can provide information on the S-M system which is new or different to that which can be obtained from 1-D tracking tasks. The primary evidence for this assertion is the larger degradation in tracking performance, on going from a 1-D to a 2-D task, observed for the Parkinsonian subjects in comparison to that observed for the normals. The value of the 2-D random pursuit tracking task has also been shown by Dalrymple Alford *et al.* (1994) where it was used to demonstrate deficits in the Parkinsonian central executive.

The second study which compared normal 1-D and 2-D tracking performance was able to demonstrate unequivocally that 2-D tracking is more difficult than 1-D tracking. By scaling the 2-D target signal so that it had the same average displacement and velocity as the 1-D signal from which it was derived, it has been possible to demonstrate that the degradation in tracking performance observed on going to 2-D occurs because of the multidimensional nature of the target signal and not simply because of an increase in average target velocity. It is hypothesised that the multidimensional nature of the 2-D tracking task places extra demands on the finite resources of the S-M system. As a consequence, the S-M system cannot use target prediction to reduce to the same extent the inherent delays within the system and also must resort to processing some of the information required to develop a movement plan in a sequential manner. This reduction in the ability to use prediction, and the sequential processing of additional information, increases the time delay in the S-M system (i.e. the time delay between when the target moves and the S-M system responds) which in turn results in an increase in tracking error. It was found that 73% of the increase in tracking error, that occurs on going from the 1-D to the 2-D task described here, could be attributed to this extra time delay. It is anticipated that the remaining 27% could be attributable to the subject performing open-loop movements to catch up to the target, a reduction in target prediction accuracy, and a reduction in movement plan accuracy. The results in §10.2.3 have shown that the size of this degradation does not increase with age and is not dependent on gender. When the 2-D target signal has a larger average velocity than the 1-D target signal, the increase in average task velocity also makes a contribution to the degradation in performance observed on going to 2-D (see §10.2.3).

On comparing performance in each coordinate of the 2-D tracking tasks to the corresponding coordinates in the 1-D tasks (see §10.2.4) it was found that the observed differences in performance were dependent on the average velocity of the 2-D task and the previous exposure of the subjects to the tasks. The dependence on average task velocity indicates that comparing performance in the common coordinates of the 2-D and 1-D tasks, as is done by Garvey and Taylor (1959), Chernikoff *et al.* (1960), Todosiev *et al.* (1967), and Ziegler (1968), is not a valid method for determining whether the S-M system requires more resources to complete a 2-D task than a 1-D task. In addition, comparing the performance in the common coordinates relies on the questionable assumption that a 2-D task is not a single task but two independent tasks performed at the same time.

The results in §9.2.6.2 show that the most accurate time delay estimation techniques for application with tracking tasks are the standard spatial or frequency domain cross-correlation technique and the least squares time delay estimation technique. It is necessary to note that before the time delay was calculated by the standard cross-correlation technique, the large dc components in the target signal and the simulated response signal were removed to avoid one of the weaknesses of this technique. This weakness is the severe distortion of the cross-correlation function, and the resulting bias of the time delay estimate towards $t = 0$, that can occur if both signals being correlated have a significant dc component. The study also demonstrated that the adapted cross-correlation technique, which was the time delay estimation technique originally used with the tracking tasks described here, provided the least accurate time delay estimates of

the five techniques compared. The inaccuracy of adapted cross-correlation arises because of the differences in the lengths of the windows used to truncate the target and response signals. The expression derived in §9.2.3, to describe the error in the time delay estimate computed by this technique, demonstrates that this difference in window length produces a linear phase error in the cross-spectrum, which shifts the position of the peak of the cross-correlation function. This phase error is large when a multispectral signal has a large number of low frequency components and the difference between the window lengths is large. If the windows are approximately the same length and the signals being correlated do not have a significant low frequency content, it is quite possible that the accuracy of the adapted cross-correlation technique could approach that of the standard technique.

11.2 SUGGESTIONS FOR FURTHER WORK

In the first clinical study (see §10.1) it was found that Parkinsonian S-M performance degraded more than that of normals on going from a 1-D to a 2-D task. It was, however, unclear whether this larger degradation in the Parkinsonian subject's performance was due to the multidimensional nature or the increased velocity/acceleration of the 2-D task. It would be worth determining whether this extra degradation in performance indicates that the Parkinsonian S-M system has more difficulty processing multidimensional information than the normal system or whether it simply reflects the general slowing down of the S-M system with Parkinson's disease. To answer this question another study of Parkinsonian S-M function would need to be conducted and performance on the 1-D tasks would have to be compared to that on the *scaled* 2-D task. If it was found that performance did degrade on going from the 1-D to the scaled 2-D task it would be possible to conclude, since both tasks have the same average displacement and velocity, that the multidimensional nature of the 2-D task was responsible for the degradation

Another area worthy of future research is the extension of the findings implicit within Eqn. (10.1), relating the error observed on a 1-D task to that observed on a 2-D task, into a full model describing how a subject tracks in a multidimensional environment. To accomplish this, it would be necessary to determine whether the processing delay observed on going from a 1-D to a 2-D task is a constant, independent of the frequency content of the target signal. It would also be necessary to develop a model describing the decrease in the accuracy of the prediction of the target's position as the number of degrees of freedom increases. It would be interesting to determine whether the extra processing delay and prediction inaccuracy can be described by the adaptive model which Neilson *et al.* (1992) has developed to describe the S-M system. The ultimate goal would be to develop a model of the S-M system which could take a target signal of any dimensionality and produce as its output a response which closely approximates that observed from the human S-M system. Ideally this model would, in a similar manner to that of Neilson *et al.* (1992), unify the known physiology of the S-M system, adaptive control theory, and the experimentally derived data describing the S-M system's characteristics. That is, the model should not simply be a mathematical model with little resemblance to the known physiological structure of the S-M system.

When performance in the vertical direction of the 2-D tasks was compared to that in the horizontal direction, it was found that in some cases the vertical performance was superior to that in the horizontal direction. Specifically, for the younger 20-40 yrs age group, vertical performance on the 2-D random pursuit tracking task was found to be superior to the horizontal performance, while with the older 55-75 yrs age group, performance in the vertical direction of the step task was found to be superior to that in the horizontal direction. It is not possible

at this stage to determine whether the above results are due to a motor superiority, a difference in tracking strategy in each direction, or even if the results observed for the two different age groups are due to the same mechanism. Further work in this area is therefore needed if the origins of the above directional biases are to be determined. This additional work may provide important information on how the S-M system changes with the ageing process.

The final area where the author sees a possibility for further work is the development of an analytic expression (similar to that given in Appendix A) that completely describes the form of the cross-correlation of two truncated sinusoids. On obtaining this expression it may then be possible to derive another expression that describes the bias, towards $t = 0$, that occurs when standard cross-correlation is used to estimate time delays. This expression would primarily be of academic interest, as the results presented in §9.2.6.2 have shown that, provided the window length is much greater than the periods of the significant high frequency components within the signals being correlated, the bias tends to be negligible (assuming insignificant dc components).

APPENDIX A

THE CROSS-CORRELATION OF TRUNCATED SINUSOIDS

In this appendix an attempt is made to derive an expression for the cross-correlation of a truncated sinusoid and a delayed version of the same sinusoid. Both the original sinusoid and the delayed version are truncated by multiplying by a rectangular window of length L. The aim is to obtain an expression confirming the conceptual viewpoint, outlined in §9.2.2.2, that the cross-correlation function can be adequately described by the product of a triangular window function, of length $2L$, and the analytic cross-correlation function of the untruncated sinusoids.

Assume that the signals to be cross-correlated are of the form

$$r(t) = w(t)\cos(2\pi u_0 t)$$

and

$$h(t) = w(t)\cos(2\pi u_0(t+\tau)),$$

where $w(t)$ is the rectangular window of length L (centered about $t = 0$), u_0 is the frequency of the sinusoids, and τ is the time delay between the two signals. An expression for the form of the cross-correlation function can most readily be obtained by transforming the problem to the frequency domain and examining the form of the cross-spectrum (see Eqn.(2.54)). Using the Fourier shift theorem (Table 2.1) and principles of convolution (Table 2.1) it is found that the cross-spectrum, $Z_{RH}(u)$, can be described by the expression

$$\begin{aligned}
Z_{RH}(u) &= \left(W(u) \odot \left(\delta(u+u_0) + \delta(u-u_0)\right)\right)^* \\
&\quad \left(W(u) \odot \left(\delta(u+u_0)\exp^{-i2\pi u_0\tau} + \delta(u-u_0)\exp^{i2\pi u_0\tau}\right)\right) \\
&= \left(W(u+u_0) + W(u-u_0)\right)^* \left(W(u+u_0)\exp^{-i2\pi u_0\tau} + W(u-u_0)\exp^{i2\pi u_0\tau}\right) \\
&= \left(|W(u+u_0)|\exp^{-iPh(W(u+u_0))} + |W(u-u_0)|\exp^{-iPh(W(u-u_0))}\right) \\
&\quad \left(|W(u+u_0)|\exp^{iPh(W(u+u_0))}\exp^{-i2\pi u_0\tau} + |W(u-u_0)|\exp^{iPh(W(u-u_0))}\exp^{i2\pi u_0\tau}\right) \\
&= |W(u+u_0)|^2\exp^{-i2\pi u_0\tau} + |W(u-u_0)|^2\exp^{i2\pi u_0\tau} \\
&\quad + |W(u+u_0)|\,|W(u-u_0)|\exp^{i(-Ph(W(u+u_0))+Ph(W(u-u_0))+2\pi u_0\tau)} \\
&\quad + |W(u-u_0)|\,|W(u+u_0)|\exp^{i(Ph(W(u+u_0))-Ph(W(u-u_0))-2\pi u_0\tau)}. \qquad \text{(A.1)}
\end{aligned}$$

On using the relationships $A\exp^{i\phi} + A\exp^{-i\phi} = 2A\cos(\phi)$ and $\cos(\phi \pm \psi) = \cos(\phi)\cos(\psi) \mp \sin(\phi)\sin(\psi)$, and observing that the first two terms in Eqn.(A.1) are in the form of a convolution, Eqn.(A.1) can be rewritten in the form

$$Z_{RH}(u) = |W(u)|^2 \odot \left(\delta(u+u_0)\exp^{-i2\pi u_0\tau} + \delta(u-u_0)\exp^{i2\pi u_0\tau}\right)$$

$$
\begin{aligned}
&+2\,|W(u-u_0)|\,|W(u+u_0)| \\
&(\cos(-Ph(W(u+u_0))+Ph(W(u-u_0)))\cos(2\pi u_0\tau) \\
&-\sin(-Ph(W(u+u_0))+Ph(W(u-u_0)))\sin(2\pi u_0\tau))\,.
\end{aligned}
$$

Noting that squaring the spectral magnitude of a rectangular function of length L produces the spectrum of a triangular function of length $2L$ (see Bracewell (1978)) and that

$$\sin(-Ph(W(u+u_0))+Ph(W(u-u_0)))=0$$

(the windows are centered about $t=0$ and therefore the phases of their spectral components must be 0 or π) allows $Z_{RH}(u)$ to be expressed in the form

$$
\begin{aligned}
Z_{RH}(u) &= W_{Tri}(u)\odot\left(\delta(u+u_0)\exp^{-i2\pi u_0\tau}+\delta(u-u_0)\exp^{i2\pi u_0\tau}\right) \\
&\quad +\cos(2\pi u_0\tau)\Upsilon(u\pm u_0), \qquad \text{(A.2)}
\end{aligned}
$$

where $W_{Tri}(u)$ is the spectrum of the triangular window function,

$$
\begin{aligned}
w_{tri}(t) &= 1-\tfrac{|t|}{L}, \quad |t|\le L \\
&= 0, \qquad\quad |t|>L,
\end{aligned}
$$

and $\Upsilon(u\pm u_0)$ represents the component of the cross-spectrum due to the cross-terms. On using the convolution theorem and taking the inverse Fourier transform of the last line in Eqn.(A.2) it can be seen that the cross-correlation of the above two sinusoids is given by the expression

$$z_{rh}(t)=w_{tri}(t)\cos(2\pi u_0(t+\tau))+\cos(2\pi u_0\tau)v(t,u_0), \qquad \text{(A.3)}$$

where $v(t,u_0)=\mathsf{F}^{-1}\{\Upsilon(u\pm u_0)\}$. Eqn.(A.3) demonstrates that the cross-correlation is indeed very similar to the product of the triangular window and the analytic cross-correlation, but that a second term is present which is dependent on u_0 and has a magnitude that is dependent on the size of the delay τ.

In lieu of an analytic expression for $v(t,u_0)$ a numerical evaluation has been performed and is shown in Fig. A.1. A mesh plot of $|v(t,u_0)|$ obtained for various values of T_0 ($T_0=1/u_0$) is plotted in Fig. A.1 and in Fig. A.2 $v(t,u_0)$ is plotted for a number of non-zero values of $\frac{L}{T_0}$ between 0 and 1. For comparision, $w_{tri}(t)$ is plotted in Fig. A.1 along the line corresponding to $u_0=0$. It can therefore be seen that, as the ratio $\frac{L}{T_0}$ becomes large, $v(t,u_0)$ becomes negligible in comparison to $w_{tri}(t)$. Thus, when L is greater than (but not necessarily much greater than) T_0, the cross-correlation can be adequately described by the first term in Eqn.(A.3). For example when $\frac{L}{T_0}=3$ the ratio of the maximum of $v(t)$ to that of $w_{tri}(t)$ is approximately 0.05. The plots in Fig. A.2 also show that when $\tau=0$ the peak of the first term in Eqn.(A.3) aligns with a local maxima of the second term and therefore the position of the maxima of $z_{rh}(t)$ must remain at $t=0$. The remaining case that needs to be considered is that where $L<T_0$ and $0<\tau<\frac{\pi}{4}$. At present the author has been unable to devise a rigorous explanation of how, in this situation, the position of the maximum of the first term in Eqn.(A.3) is influenced by the second term and this case therefore requires further work. It is worth noting, however, that in the multispectral case the cross-correlation should be adequately described by the product of a triangular window and the analytic cross-correlation, if the signals being correlated have high frequency components of significant magnitude. The reason for this is that the sum of all the terms due to the cross-terms in the cross-spectrum should be much less than the sum of all the terms that are scaled by $w_{tri}(t)$.

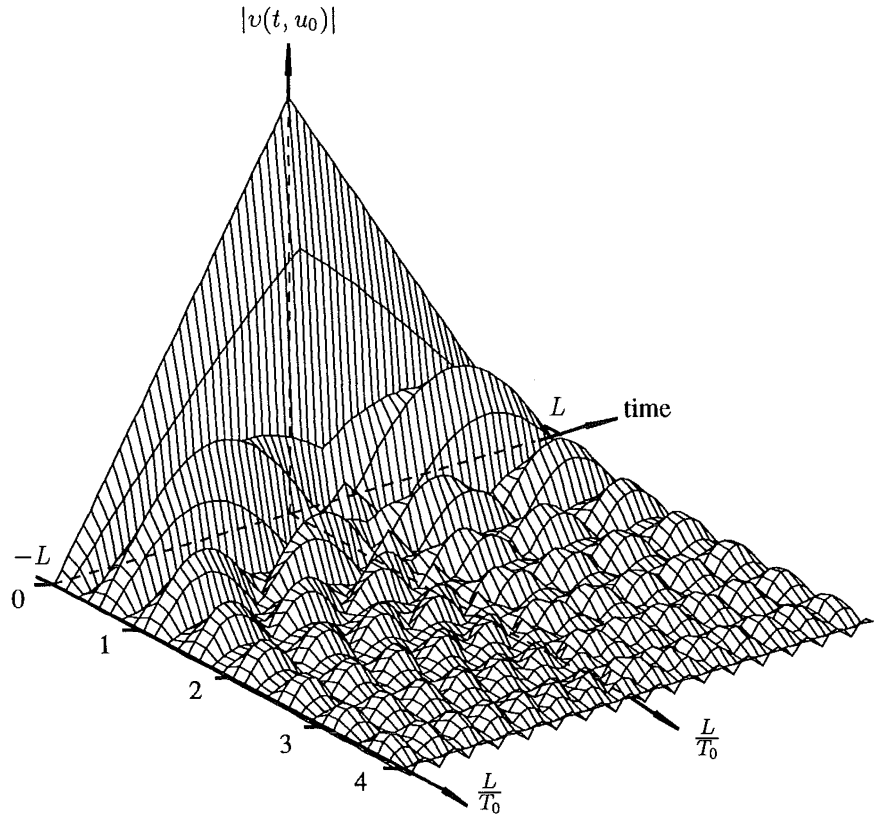

Figure A.1 The magnitude of $v(t, u_0)$ as a function of the ratio of the window length L to the period T_0 ($T_0 = \frac{1}{u_0}$) of the sinusoids being correlated.

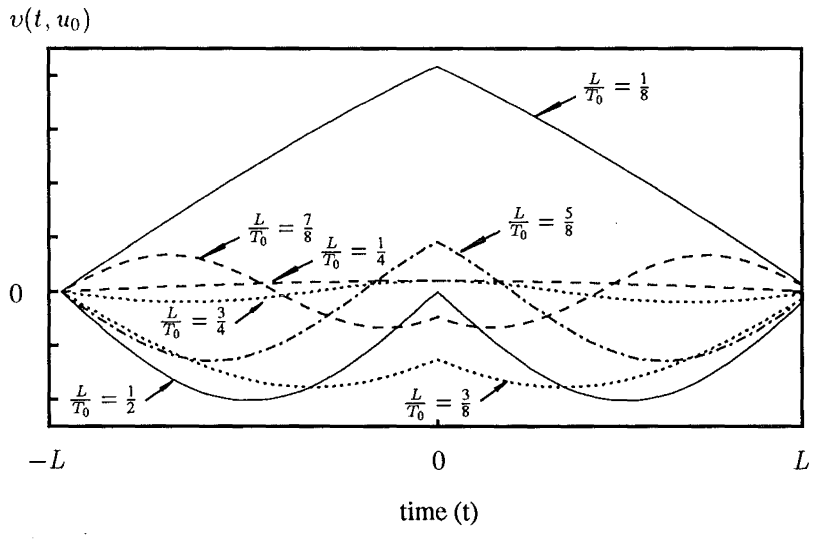

Figure A.2 The form of $v(t, u_0)$, as a function of the ratio of the window length L to the period T_0 ($T_0 = \frac{1}{u_0}$) of the sinusoids being correlated, for various values of $\frac{L}{T_0}$ between 0 and 1.

REFERENCES

ABDEL-MALEK, A., MARKHAM, C.H., MARMARELIS, P.Z. and MARMARELIS, V.Z. (1988), 'Quantifying deficiencies associated with Parkinson's disease by use of time-series analysis', *Electroencephalogr. Clin. Neurophysiol.*, Vol. 69, pp. 24–33.

ADAMS, J.A. (1952), 'Warm-up decrement in performance on the pursuit rotor', *Amer. J. Psychol.*, Vol. 65, pp. 404–414.

AGNEW, P.J. and MAAS, F. (1982), 'Hand function related to age and sex', *Arch. Phys. Med. and Rehab.*, Vol. 63, pp. 269–271.

ANDREWS, H.C. and HUNT, B.R. (1977), *Digital image restoration*, Prentice-Hall, New Jersey.

ANTON, H. (1984), *Elementary linear algebra*, John Wiley and Sons, New York, 4th ed.

ARCHER, E.J., KENT, G.W. and MOTE, F.A. (1956), 'Effect of long-term practice and time-on-target information feedback on a complex tracking task', *J. Exp. Psychol.*, Vol. 51, pp. 103–112.

AVITZOUR, D. (1991), 'Time delay estimation at high signal-to-noise ratios', *IEEE Trans. Aerosp. Electron. Sys.*, Vol. 27, pp. 234–237.

BADDELEY, A., LOGIE, R., BRESSI, S., DELLA SALA, S. and SPINNLER, H. (1986), 'Dementia and working memory', *Quart. J. Exp. Psychol.*, Vol. 38A, pp. 603–618.

BARRETT, A.H. and MYERS, P.C. (1975), 'Subcutaneous temperatures: a method of noninvasive sensing', *Science*, Vol. 190, pp. 669–671.

BATES, R.H.T. (1976), 'A stochastic image restoration procedure', *Optics Comm.*, Vol. 19, No. 2, pp. 240–244.

BATES, R.H.T. and MCDONNELL, M.J. (1986), *Image restoration and reconstruction*, Clarendon Press.

BATES, R.H.T. and MNYAMA, D. (1986), 'The status of practical Fourier phase retrieval', In *Advances in electronics and electron physics*, Academic Press, London, pp. 1–64.

BATES, R.H.T. and NAPIER, P.J. (1971), 'A suggestion for determining antenna pattern phase from holographic type measurements', *Proc. Inst. Radio Electron. Eng. Aust.*, April, pp. 164–166.

BATES, R.H.T., NAPIER, P.J., MCKINNON, A.E. and MCDONNELL, M.J. (1976), 'Self-consistent deconvolution: 1 theory', *Optik*, Vol. 44, pp. 183–201.

BATES, R.H.T., GARDEN, K.L. and PETERS, T.M. (1983), 'Overview of computerised tomography with emphasis on future developments', *Proc. IEEE*, Vol. 71, No. 3, pp. 356–372.

BATES, R.H.T., QUEK, B.K. and PARKER, C.R. (1990), 'Some implications of zero sheets for blind deconvolution and phase retrieval', *J. Opt. Soc. Am. A*, Vol. 7, No. 3, pp. 468–479.

BATES, R.H.T., SMITH, V.A. and MURCH, R.D. (1991), 'Manageable multidimensional inverse scattering theory', *Physics Reports*, Vol. 201, No. 4, pp. 187–277.

BEARCE, P.G. and MYERS, J.L. (1990), *Principles and practice of adult health nursing*, The CV Mosby Company, Toronto.

BEAUCAMP, K.G. (1973), *Signal processing using analog and digital signals*, Vol. 1973, George Allen and Unwin Ltd, London.

BEDINI, L., FOSSI, S. and REGGIANNINI, R. (1986), 'Generalised crosscorrelator with data-estimated weighting function: a simulation analysis', *IEE Proc., Part F*, Vol. 133, pp. 195–198.

BEHBEHANI, K., KONDRASKE, G.V. and RICHMOND, J.R. (1988), 'Investigation of upper extremity visuomotor control performance measures', *IEEE Trans. Biomed. Eng.*, Vol. 35, pp. 518–525.

BENDAT, J.S. and PIERSLO, A.G. (1980), *Engineering applications of correlation and spectral analysis*, Wiley, New York.

BIEMOND, J., VAN DER PUTTEN, F.G. and WOODS, J.W. (1988), 'Identification and restoration of images with symmetric noncausal blurs', *IEEE Trans Circuits Syst.*, Vol. 35, pp. 385–393.

BIEMOND, J., LAGENDIJK, R.L. and MERSEREAU, R.M. (1990), 'Iterative methods for image deblurring', *Proc. IEEE*, Vol. 78, pp. 856–883.

BLITZ, J. (1991), *Electrical and magnetic methods of nondestructive testing*, IOP Publishing, Bristol, England.

BLOXHAM, C.A., MINDEL, T.A. and FRITH, C.D. (1984), 'Initiation and execution of predictable and unpredictable movements in Parkinson's disease', *Brain*, Vol. 107, pp. 371–384.

BONES, P.J., SATHERLEY, B.L. and WATSON, R.W. (1992), 'Direct deconvolution of noisy blurred images', In *IEE 4th Conference on Image Processing and its Applications*, pp. 353–356.

BONES, P.J., SATHERLEY, B.L., PARKER, C. and WATSON, R.W. (1993), 'Non-iterative methods for image deconvolution', In IDELL, P.S. (Ed.), *Digital image recovery and synthesis II, vol 2029*, SPIE, pp. 2–13.

BORN, M. and WOLF, E. (1970), *Principles of optics*, Pergamon Press, London, 4th ed.

BOTWINICK, J. (1981), 'Neuropsychology of aging', In FILSKOR, S.B. and BOLL, T. (Eds.), *In Handbook of Clinical Neuropsychology*, Wiley, New York.

BRACEWELL, R.N. (1978), *The Fourier transform and its applications*, McGraw Hill, Tokyo, 2nd ed.

BRIGHAM, E.O. (1988), *The fast Fourier transform and its applications*, Prentice Hall, Englewood Cliffs NJ.

BRUCK, Y.M. and SODIN, L.G. (1979), 'On the ambiguity of the image reconstruction problem', *Optics Comm.*, Vol. 30, pp. 304–308.

BURDEN, R.L. and FAIRES, J.D. (1985), *Numerical analysis*, Prindle, Weber and Schmidt, Boston, 3rd ed.

CAMERON, J.R. and SKOFRONICK, J.G. (1978), *Medical physics*, John Wiley and Sons, New York.

CARLSON, N.R. (1986), *Physiology of behaviour*, Allyn and Bacon, Massachusetts.

CASSELL, K.J. (1973), 'The usefulness of a temporal correlation technique in the assessment of human motor performance on a tracking device', *Med. Biol. Eng.*, pp. 755–761.

CASSELL, K., SHAW, K. and STERN, G. (1973), 'A computerized tracking technique for the assessment of Parkinsonian motor disabilities', *Brain*, Vol. 96, pp. 815–826.

CASTLEMAN, K.R. (1979), *Digital Image Processing*, Prentice-Hall, New Jersey.

CHERNIKOFF, R., DUEY, J.W. and TAYLOR, F.V. (1960), 'Two-dimensional tracking with identical and different control dynamics in each coordinate', *J. Exp. Psychol*, Vol. 60, pp. 318–322.

CIVANLAR, M.R. and TRUSSELL, H.J. (1986), 'Digital signal restoration using fuzzy sets', *IEEE Trans Acoustic Speech Signal Process.*, Vol. 34, pp. 919–936.

CONNER, F.R. (1989), *Antennas*, Edward Arnold, London, 2nd ed.

COOPER, R., MCCALLUM, W.C. and CORNWAITE, S.P. (1989), 'Slow potential changes related to the velocity of target movement in a tracking task', *Electroencephalogr. Clin. Neurophysiol.*, Vol. 72, pp. 232–239.

CORKIN, S. (1968), 'Acquisition of motor skill after medial temporal-lobe excision', *Neuropsychologica*, Vol. 6, pp. 255–265.

CRAIK, K.J.W. (1948), 'Theory of the human operator in control systems', *Brit. J. Psychol.*, Vol. 38, pp. 142–148.

CROFT, D. and JONES, R.D. (1987), 'The value of off-road tests in the assessment of driving potential of unlicensed disabled people', *Brit. J. Occup. Ther.*, Vol. 50, pp. 357–361.

CURTIS, S.R. and OPPENHEIM, A.V. (1987), 'Reconstruction of multidimensional signals from zero crossings', *J. Opt. Soc. Am. A*, Vol. 4, pp. 221–231.

DALRYMPLE-ALFORD, J.C., KALDERS, A.S., JONES, R.D. and WATSON, R.W. (1994), 'A central executive deficit in patients with Parkinson's disease', *J. Neurol. Neurosurg. Psychiatry, In Press*, Vol. 57.

DAMOS, D.L. and WICKENS, C.D. (1977), 'Dual-task performance and the Hick-Hyman law of choice reaction time', *J. Motor Behav.*, Vol. 9, pp. 209–217.

DAVEY, B.L.K. (1989), *Advances in blind deconvolution*, PhD thesis, University of Canterbury, Christchurch, New Zealand.

DAVEY, B.L.K., SINTON, A.M. and BATES, R.H.T. (1986), 'Zero-and-add', *Opt. Eng.*, Vol. 25, No. 6, pp. 765–771.

DAY, B.L., DICK, J.P.R. and MARSDEN, C.D. (1984), 'Patients with Parkinson's disease can employ a predictive motor strategy', *J. Neurol. Neurosurg. Psychiatry*, Vol. 47, pp. 1299–1306.

DE SOUZA, L.H., LANGTON HEWER, R., LYNN, P.A., MILLER, S. and REED, G.A.L. (1980), 'Assessment of recovery of arm control in hemiplegic stroke patient. 2. Comparison of arm function tests and pursuit tracking in relation to clinical recovery.', *Int. Rehab. Med.*, Vol. 2, pp. 10–16.

DIXON, W.J. (Ed.) (1981), *BMDP Statistical Software 1981*, University of California Press, Berkeley.

DONGARRA, J.J., MOLER, C.B., BUNCH, J.R. and STEWART, G.W. (1979), *LINPACK Users' Guide*, SIAM.

DUBROVIN, B.A., FOMENKO, A.T. and NOVIKOV, S.P. (1985), *Modern geometry: methods and applications*, Springer-Verlag, New York.

DUFFY, F.H., IYER, V.G. and SURWILLO, W.W. (1989), *Clinical electroencephalography and topographic brain mapping*, Springer Verlag, New York.

ELLIOT, R.S. (1966), *General electromagnetic theory*, McGraw Hill, New York.

ENRIGHT, S.A. (1992), *Towards quantitative computed tomography*, PhD thesis, University of Canterbury, Christchurch, New Zealand.

ERIKESEN, C.W., HAMLIN, R.M. and BREITMEYER, R.G. (1970), 'Temporal factors in visual perception as related to aging', *Percept. Psychophys.*, Vol. 7, pp. 354–356.

EVES, H. (1965), *A survey of geometry*, Allyn and Bacon Inc, Boston.

EYÜBOĞLU, B.M., BROWN, B.H. and BARBER, D.C. (1989), 'In vivo imaging of cardiac related impedance changes', *IEEE Eng. Med. Biol. Soc. Magazine*, Vol. 8, pp. 39–45.

EYZAGUIRRE, C. and FIDONE, S.J. (1975), *Physiology of the nervous system*, Year Book Medical Publishers Incorporated, Chicago.

FERTNER, A. and SJÖLUND, A. (1986), 'Comparison of various time delay estimation methods by computer simulation', *IEEE Trans Acoustics Speech Signal Process.*, Vol. 34, pp. 1329–1330.

FIENUP, J.R. (1979), 'Space object imaging through a turbulent atmosphere', *Opt. Eng.*, Vol. 18, pp. 529–534.

FIENUP, J.R. (1982), 'Phase retrieval algorithms: a comparison', *Appl. Optics*, Vol. 21, pp. 2758–2769.

FIENUP, J.R. and WACKERMAN, C.C. (1986), 'Phase-retrieval stagnation problems and solutions', *J. Opt. Soc. Am. A*, Vol. 3, pp. 1897–1907.

FLEISHMAN, E.A. (1954), 'Dimensional analysis of psychomotor abilities', *J. Exp. Psychol.*, Vol. 48, pp. 437–454.

FLEISHMAN, E.A. (1958), 'Dimensional analysis of movement reactions', *J. Exp. Psychol.*, Vol. 5, pp. 438–453.

FLEISHMAN, E.A. (1972), 'Structure and measurement of psychomotor abilities', In SINGER, R.N. (Ed.), *The psychomotor domain: movement behaviour*, Lea and Febiger, Philadelphia.

FLEISHMAN, E.A. and ELLISON, G.D. (1962), 'A factor analysis of fine manipulative tests', *J. Appl. Psychol.*, Vol. 46, pp. 96–105.

FLOWERS, K.A. (1976), 'Visual 'closed-loop' and 'open-loop' characteristics of voluntary movement in patients with Parkinsonism and intention tremor', *Brain*, Vol. 99, pp. 269–310.

FLOWERS, K.A. (1978), 'Lack of prediction in the motor behaviour of Parkinsonism', *Brain*, Vol. 101, pp. 35–52.

FRACKER, M.L. and WICKENS, C.D. (1989), 'Resources, confusions, and compatibility in dual-axis tracking: displays, controls, and dynamics', *J. Exp. Psychol.*, Vol. 15, pp. 80–96.

FRASER, C. and BENTON, J. (1983), 'A study of adult hand strength', *Occup. Ther.*, Vol. 46, pp. 296–299.

FRIGHT, W.R. (1984), *The Fourier phase problem*, PhD thesis, University of Canterbury, Christchurch, New Zealand.

FRITH, C.D. and LANG, R.J. (1979), 'Learning and reminiscence as a function of target predictability in a two-dimensional tracking task', *Quart. J. Exp. Psychol.*, Vol. 31, pp. 103–109.

FRITH, C.D., BLOXHAM, C.A. and CARPENTER, K.N. (1986), 'Impairments in the learning and performance of a new manual skill in patients with Parkinson's disease', *J. Neurol. Neurosurg. Psychiatry*, Vol. 49, pp. 661–668.

FUCHS, B.A. and LEVIN, V.I. (1961), *Functions of a complex variable and some of their applications*, Vol. 2, Pergamon Press, London.

GARDENIER, P.H. (1990), *Antenna aperture phase retrieval*, PhD thesis, University of Canterbury, Christchurch, New Zealand.

GARVEY, W.D. and TAYLOR, F.V. (1959), 'Interactions among operator variables, system dynamics, and task induced stress', *J. Appl. Psychol.*, Vol. 43, pp. 79–85.

GENNERY, D.B. (1973), 'Determination of optical transfer function by inspection of the frequency-domain plot', *J. Opt. Soc. Am. A*, Vol. 63, pp. 1571–1577.

GERCHBERG, R.W. and SAXTON, W.O. (1972), 'A practical algorithm for the determination of phase from image and diffraction plane pictures', *Optik*, Vol. 35, pp. 237–246.

GHIGLIA, D.C., ROMERO, L.A. and MASTIN, G.A. (1993), 'Systematic approach to two-dimensional blind deconvolution by zero sheet separation', *J. Opt. Soc. Am. A*, Vol. 10, pp. 1024–1036.

GOODMAN, J.W. (1968), *Introduction to Fourier optics*, McGraw Hill, New York.

GOPHER, D. and NAVON, D. (1980), 'How is performance limited: testing the notion of central capacity', *Acta Psychol.*, Vol. 46, pp. 161–180.

GULL, S.F. and SKILLING, J. (1984), 'Maximum entropy method in image processing', *IEE Proc. Part F*, Vol. 131.

GUSTAVSSON, M., IVANSSON, S., MORÉN, P. and PHIL, J. (1986), 'Seismic borehole tomography - measurement system and field studies', *Proc. IEEE*, Vol. 74, pp. 339–346.

HAYES, M.H. and MCCLELLAN, J.H. (1982), 'Reducible polynomials in more than one variable', *Proc. IEEE*, Vol. 70, pp. 197–198.

HAYKIN, S. (1983), *Communication systems*, John Wiley and Sons, New York, 2nd ed.

HELSON, H. (1949), 'Design of equipment and optimal human operation', *Amer. J. Psychol.*, Vol. 62, pp. 473–497.

HINSHAW, W.S. and LENT, A.H. (1983), 'An introduction to NMR imaging: from the Bloch equation to the imaging equation', *Proc. IEEE*, Vol. 7, pp. 338–350.

HOEHN, M.M. and YAHR, M.D. (1967), 'Parkinsonism: onset, progression, and mortality', *Neurology*, Vol. 17, pp. 427–442.

HULL, B. and VERNON, J. (1988), *Non-destructive testing*, MacMillan Education Ltd, London.

HUNTER, I.W., KEARNEY, R.E. and JONES, L.A. (1987), 'Estimation of the conduction velocity of muscle action potentials using phase and impulse response function techniques', *Med. Biol. Eng. Comput.*, Vol. 25, pp. 121–126.

HURT, N.E. (1989), *Phase retrieval and zero crossings: mathematical methods in image reconstruction*, Kluwer Academic Publishers, Dordrecht, The Netherlands.

HYDE, V. (1992), *Godzone Skies: Astronomy for New Zealanders*, University of Canterbury Press, Christchurch.

ISON, K.T. (1988), 'Assessment of control ability in the physically handicapped', *J. Biomed. Eng.*, Vol. 10, pp. 430–437.

IZRAELEVITZ, D. and LIM, J.S. (1987), 'A new direct algorithm for image reconstruction from Fourier transform magnitude', *IEEE Trans. Acoustics Speech Signal Process.*, Vol. 35, No. 4, pp. 511–519.

JACKSON, K.F. (1958), 'Behaviour in controlling a combination of systems', *Ergonomics*, Vol. 2, pp. 52–62.

JACOVITTI, G. and SCARANO, G. (1993), 'Discrete time techniques for time delay estimation', *IEEE Trans. Signal Process.*, Vol. 41, pp. 525–533.

JASZCZAK, R.J. (1988), 'Tomographic radiopharmaceutical imaging', *Proc. IEEE*, Vol. 76, pp. 1079 – 1094.

JENKINS, M.A. and TRAUB, J.F. (1970), 'A three-stage variable-shift iteration for polynomial zeros and its relation to generalized Rayleigh iteration', *Numer. Math.*, Vol. 14, pp. 252–263.

JENKINS, M.A. and TRAUB, J.F. (1972), 'Zeros of a complex polynomial', *Comm. ACM*, Vol. 15, No. 2, pp. 97–99.

JONES, D.S. (1986a), *Acoustic and electromagnetic waves*, Clarendon Press, Oxford.

JONES, R.D. (1986b), *Measurement of normal and abnormal sensory-motor function by a computerized test battery*, PhD thesis, Otago University, Dunedin, New Zealand.

JONES, R.D. and DONALDSON, I.M. (1981), 'Measurement of integrated sensory-motor function following brain damage by a preview tracking task', *Int. Rehabil. Med.*, Vol. 3, pp. 71–83.

JONES, R.D. and DONALDSON, I.M. (1986), 'Measurement of sensory-motor integrated function in neurological disorders: three computerized tracking tasks', *Med. Biol. Eng. Comput.*, Vol. 24, pp. 536–540.

JONES, R.D. and DONALDSON, I.M. (1989), 'Tracking tasks and the study of predictive motor planning in Parkinson's disease', In *Proc. 11th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 1055–1056.

JONES, R.D. and WATSON, R.W. (1993), 'A two-dimensional step tracking task and its application to Parkinson's disease', In *Proc. 15th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, San Diego, pp. 1416–1417.

JONES, R.D., GIDDENS, H. and GROFT, D. (1983), 'Assessment and training of brain damaged drivers', *Amer. J. Occup. Ther.*, Vol. 37, pp. 754–760.

JONES, R.D., WILLIAMS, L.R.T. and WELLS, J.E. (1986), 'Effects of laterality, sex, and age on computerized sensory-motor tests', *J. Hum. Motor Studies*, Vol. 12, pp. 163–182.

JONES, R.D., DONALDSON, I.M. and PARKIN, P.J. (1989), 'Impairment and recovery of ipsilateral sensory-motor function following unilateral cerebral infarction', *Brain*, Vol. 112, pp. 113–132.

JONES, R.D., DONALDSON, I.M., PARKIN, P.J. and COPPAGE, S.A. (1990), 'Impairment and recovery profiles of sensory-motor function following stroke: single-case graphical analysis techniques', *Int. Disabil. Studies*, Vol. 12, pp. 141–148.

JONES, R.D., DONALDSON, I.M. and TIMMINGS, P.L. (1992), 'Impairment of high-contrast visual acuity in Parkinson's disease', *Mov. Disord.*, Vol. 7, pp. 232–238.

JONES, R., SHARMAN, N., WATSON, R. and MUIR, S. (1993), 'A PC-based battery of tests for quantitative assessment of upper-limb sensory-motor function in brain disorders', In *Proc. 15th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, San Diego, pp. 1414–1415.

KATSAGGELOS, A.K. (1989), 'Iterative image restoration algorithms', *Opt. Eng.*, Vol. 28, pp. 735–748.

KATSAGGELOS, A.K. (1991), *Digital image restoration*, Springer-Verlag, Berlin.

KNOLL, G.F. (1983), 'Single-photon emission computed tomography', *Proc. IEEE*, Vol. 71, pp. 320–329.

KNOX, K.T. and THOMPSON, B.J. (1974), 'Recovery of images from atmospherically degraded short exposure photographs', *Astrophys. J. Lett.*, Vol. 193, pp. 45–48.

KONDRASKE, G.V., POTVIN, A.R., TOURTELLOTTE, W.W. and SYNDULKO, K. (1984), 'A computer-based system for automated quantitation of neurologic function', *IEEE Trans. Biomed. Eng.*, Vol. 31, pp. 401–414.

KRAUS, J.D. (1966), *Radio astronomy*, McGraw Hill, New York.

KREYSZIG, E. (1983), *Advanced engineering mathematics*, John Wiley and Sons, New York, 5th ed.

LABEYRIE, A. (1970), 'Attainment of diffraction-limited resolution in large telescopes by Fourier analysing speckle patterns in star images', *Astronomy and Astrophysics*, Vol. 6, No. 1, pp. 85–87.

LANE, R.G. (1988), *Blind deconvolution and phase retrieval*, PhD thesis, University of Canterbury, Christchurch, New Zealand.

LANE, R.G. and BATES, R.H.T. (1987), 'Automatic multidimensional deconvolution', *J. Opt. Soc. Am. A*, Vol. 4, No. 1, pp. 180–188.

LANE, R.G., FRIGHT, W.R. and BATES, R.H.T. (1987), 'Direct phase retrieval', *IEEE Trans. Acoustics Speech Signal Process.*, Vol. 35, No. 4, pp. 520–525.

LANGE, F.H. (1967), *Correlation techniques: foundations and applications of correlation analysis in modern communications, measurement and control*, Iliffe Books Ltd, London.

LAWSON, C.L. and HANSON, R.J. (1974), *Solving least squares problems*, Prentice Hall, Englewood Cliffs, New Jersey.

LE VINE, D.M., WILHEIT, T.T., MURPHY, R.E. and SWIFT, C.T. (1989), 'A multifrequency microwave radiometer of the future', *IEEE Trans. Geosci. Remote Sensing*, Vol. 27, pp. 193–199.

LEE, H. (1990), 'Review of image-blur in a photographic system using the principles of optics', *Opt. Eng.*, Vol. 29, pp. 405–421.

LEVI, A. and STARK, H. (1987), 'Restoration from phase and magnitude by generalised projections', In STARK, H. (Ed.), *Image recovery theory and applications*, Academic Press, London.

LOHMANN, A., WEIGELT, G. and WIRNITZER, B. (1983), 'Speckle masking in astronomy: triple correlation theory and applications', *Appl. Opt.*, Vol. 22, pp. 4028–4037.

LUDEMAN, L.C. (1987), *Fundamentals of digital signal processing*, John Wiley and Sons, New York.

LYNN, P.A., REED, G.A.L., PARKER, W.R. and HEWER, R.L. (1977), 'Some applications of human operator research to the assessment of disability in stroke', *Med. Biol. Eng.*, Vol. 15, pp. 184–188.

LYNN, P.A., PARKER, W.R., REED, G.A.L. and PILSWORTH, B.W. (1979), 'New approaches to modelling the disabled human operator', *Med. Biol. Eng. Comput.*, Vol. 17, pp. 344–348.

MCCANN, H.A., SHARP, J.C., KINTER, T.M., MCEWAN, C.N., BARILLOT, C. and GREENLEAF, J.F. (1988), 'Multidimensional ultrasonic imaging for cardiology', *Proc. IEEE*, Vol. 76, pp. 1063–1071.

MCCORMACK, J.E., JUNKIN, G. and ANDERSON, A.P. (1990), 'Microwave metrology of reflector antennas from a single amplitude', *IEE Proc. Part H*, Vol. 137, pp. 276–284.

MCDONNELL, M.J. and BATES, R.H.T. (1975), 'Preprocessing of degraded images to augment existing restoration methods', *Comput. Graphics Image Process.*, Vol. 4, pp. 25–39.

MILLANE, R.P. (1990), 'Phase retrieval in crystallography and optics', *J. Opt. Soc. Am. A*, Vol. 7, pp. 394–411.

MODDEMEIJER, R. (1991), 'On the determination of the position of extrema of sampled correlators', *IEEE Trans. Signal Process.*, Vol. 39, No. 1, pp. 216–219.

MUIR, S.R., JONES, R.D., ANDREAE, J.H. and DONALDSON, I.M. (1992), 'Parkinson's disease - rhythm and finger tapping blues', In *Proc. Ann. Conf. of N.Z. Branch of Australasian College of Phys. Sci. Eng. Med., (Abstract)*, pp. 13–14.

NAGRATH, I.J. and GOPAL, M. (1982), *Control systems engineering*, Wiley Eastern Ltd, New Delhi, 2nd ed.

NAPIER, P.J. and BATES, R.H.T. (1973), 'Antenna - aperture distributions from holographic type of radiation-pattern measurement', *Proc. IEE*, Vol. 120, pp. 30–34.

NAPIER, P.J. and BATES, R.H.T. (1974), 'Inferring phase information from modulus information in two-dimensional aperture synthesis', *Astron. Astrophys. Suppl. Ser.*, Vol. 15, pp. 427–430.

NASH, J., NEILSON, P.D. and O'DWYER, N.J. (1989), 'Reducing spasticity to control muscle contracture of children with cerebral palsy', *Dev. Med. Child Neurol.*, Vol. 31, pp. 471–480.

NAVON, D., GOPHER, D., CHILLAG, N. and SPITZ, G. (1984), 'On separability of and interference between tracking dimensions in dual-axis tracking', *J. Motor Behav.*, Vol. 16, pp. 364–391.

NEILSON, P.D., O'DWYER, N.J. and NEILSON, M.D. (1988), 'Stochastic prediction in pursuit tracking: An experimental test of adaptive model theory', *Biol. Cybernetics*, Vol. 58, pp. 113–122.

NEILSON, P.D., O'DWYER, N.J. and NASH, J. (1990), 'Control of isometric muscle activity in cerebral palsy', *Dev. Med. Child Neurol.*, Vol. 32, pp. 778–788.

NEILSON, P.D., NEILSON, M.D. and O'DWYER, N.J. (1992), 'Adaptive model theory: application to disorders of motor control', In SUMMERS, J.J. (Ed.), *Approaches to the study of motor control and learning*, North-Holland, Amsterdam.

NEWMAN, B.B. and HILDEBRANDT, J. (1987), 'Blind image restoration', *The Australian Computer Journal*, Vol. 19, pp. 126–133.

NIETO-VESPERINAS, M. and DAINTY, J.C. (1986), 'Phase recovery for two-dimensional images by polynomial factorisation', *Optics Comm.*, Vol. 58, pp. 83–88.

NIETO-VESPERINAS, M. and MENDEZ, J.A. (1986), 'Phase retrieval by Monte-Carlo methods', *Optics Comm.*, Vol. 59, pp. 83–88.

NIETO-VESPERINAS, M., NAVARRO, R. and FUENTES, F.J. (1988), 'Performance of simulated annealing algorithm for phase retrieval', *J. Opt. Soc. Am. A*, Vol. 5, pp. 30–38.

OPPENHEIM, A.V. and SCHAFER, R.W. (1975), *Digital signal processing*, Prentice-Hall, Englewood Cliffs, New Jersey.

PARKER, C.R. (1992), 'Image deblurring via zero sheets', In BONES, P. (Ed.), *Proceedings of the 6th N.Z. Image Processing Workshop*, pp. 5–10.

PARKER, C.R. (1994), *Advances in zero-based phase retrieval (In preparation)*, PhD thesis, University of Canterbury, Christchurch, New Zealand.

PARKER, C.R. and BONES, P.J. (1992), 'Convergence of iterative phase retrieval improved by utilizing zero sheets', *Optics Comm.*, Vol. 92, pp. 209–214.

PERKIN, G.D. (1986), *Basic neurology*, Ellis Horwood Ltd, Chichester, England.

PERSOON, E. and KING-SUN, F. (1977), 'Shape discrimination using Fourier descriptors', *IEEE Trans. Systems, Man, Cybern.*, Vol. 7, pp. 170–179.

POTVIN, A.R. and TOURTELLOTTE, W.W. (1975), 'The neurological examination: advancements in its quantification', *Arch. Phys. Med. Rehabil.*, Vol. 56, pp. 425–437.

POTVIN, A.R., TOURTELLOTTE, W.W., PEW, R.W., ALBERS, J.W., HENDERSON, W.G. and SYNDER, D.N. (1973), 'The importance of age effects on performance in the assessment of clinical trails', *J. Chronic Dis.*, Vol. 26, pp. 699–717.

POULTON, E.C. (1952), 'Perceptual anticipation in tracking with two-pointer and one-pointer displays', *Brit. J. Psychol.*, Vol. 43, pp. 222–229.

POULTON, E.C. (1974), *Tracking skill and manual control*, Academic Press, New York.

PRESS, W.H., FLANNERY, B.R., TEUKOLSKY, S.A. and VETTERLING, W.T. (1986), *Numerical recipes*, Cambridge University Press, New York.

PULLMAN, S.L., WATTS, R.L., JUNCOS, J.L., CHASE, T.N. and SANES, J.N. (1988), 'Dopaminergic effects on simple and choice reaction time performance in Parkinson's disease', *Neurology*, Vol. 38, pp. 249–254.

PYKETT, I.L. (1982), 'NMR imaging in medicine', *Scient. Amer.*, May, pp. 54–64.

RABABY, N., KEARNEY, R.E. and HUNTER, I.W. (1989), 'Method for EMG conduction velocity estimation which accounts for input and output noise', *Med. Biol. Eng. Comput.*, Vol. 27, pp. 125–129.

RADER, C.M. and STEINHARDT, A.O. (1986), 'Hyperbolic householder transformations', *IEEE Trans. Acoustics Speech Signal Process.*, Vol. 34, pp. 1589–1602.

RADER, C.M. and STEINHARDT, A.O. (1988), *Linear algebra in signals systems and control*, SIAM, Philadelpha.

RALSTON, A. and RABINOWITZ, P. (1978), *A first course in numerical analysis*, McGraw-Hill, New York.

RAMO, S., WHINNERY, J.R. and VAN DUZER, T. (1984), *Fields and waves in communication electronics*, J Wiley and Sons, New York, 2nd ed.

REQUICHA, A.A.G. (1980), 'The zeros of entire functions: Theory and engineering applications', *Proc. IEEE*, Vol. 68, No. 3, pp. 308–328.

RILEY, P.O. and ROSEN, M.J. (1987), 'Evaluating manual control devices for those with tremor disability', *J. Rehabil. Res. Dev.*, Vol. 24, pp. 99–110.

ROBB, R.A. (1985), *Three-dimensional biomedical imaging*, CRC Press, Florida.

ROCHOW, T.G. and ROCHOW, E.G. (1980), *An introduction to microscopy by means of light, electrons, X-rays or ultrasound*, Plenum Press, London, 3rd ed.

ROHLFS, K. (1986), *Tools of radio astronomy*, Springer Verlag, New York.

ROSENFELD, A. and KAK, A.C. (1982), *Digital picture processing*, Academic Press, New York, 2nd ed.

ROSENTHAL, M., GRIFFITH, E.R., BOND, M.R. and MILLER, J.D. (1983), *Rehabilitation of the head injured adult*, F A Davis Company, Philadelphia.

ROY, A.E. and CLARKE, D. (1982), *Astronomy: principles and practice*, Adam Hilger Ltd, Bristol, 2nd ed.

SABINS, F.F. (1978), *Remote sensing principles and interpretation*, W.H. Freeman, San Francisco.

SATHERLEY, B.L. (1994), *Zero-based 2-D ensemble deconvolution and E.E.G. spectral topography (In preparation)*, PhD thesis, University of Canterbury, Christchurch, New Zealand.

SATHERLEY, B.L. and BONES, P.J. (1993), 'Using zero tracks to blindly deconvolve blurred ensembles', *Appl. Opt. (In Press)*.

SATHERLEY, B.L. and PARKER, C.R. (1993), 'Two-dimensional image reconstruction from zero sheets', *Optics Lett.*, Vol. 18, pp. 2053–2055.

SATHERLEY, B.L., PARKER, C.P. and BONES, P.J. (1994), 'Image reconstruction from zeros of the z-transform', In *Proc. Int. Conf. Acoustics Speech Signal Process.*, pp. 465–468.

SCARBROUGH, K., AHMED, N. and CARTER, C. (1981), 'On the simulation of a class of time delay estimation algorithms', *IEEE Trans. Acoustics Speech Signal Process.*, Vol. 29, pp. 534–539.

SCHMIDT, R. (1988), *Motor control and learning*, Human Kinetics Publishers, Illinois, 2nd ed.

SCHNEIDER, J.S. and LIDSKY, T.I. (Eds.) (1987), *Basal ganglia and behaviour: sensory aspects of motor functioning*, Han Huber, Toronto.

SCIVIER, M.S. and FIDDY, M.A. (1985), 'Phase ambiguities and the zeros of multidimensional band-limited functions', *J. Opt. Soc. Am. A*, Vol. 2, No. 5, pp. 693–697.

SCORER, R.S. (1986), *Cloud investigation by satellite*, Halsted Press, New York.

SCRASE, A.N., BONES, P.J., SATHERLEY, B.L. and WATSON, R.W. (1992), 'Visualisation of the image zero sheet', In *Proceedings of the 7th N.Z. Image Processing Workshop*, pp. 107–112.

SEARS, F.W., ZEMANSKY, M.W. and YOUNG, H.D. (1980), *University physics*, Addison Wesley, Reading, Massachusetts, 5th ed.

SEZAN, M.I. and STARK, H. (1982), 'Image reconstruction by the method of convex projections: Part 2 - applications and numerical results', *IEEE Trans. Med. Imaging*, Vol. 1, pp. 95–101.

SHEIKH, K., SMITH, D.S., MEADE, T.W., GOLDENBURG, E., BRENNAN, P.J. and KINSELLA, G. (1979), 'Repeatability and validity of a modified activities of daily living (ADL) index.', *Int. Rehab. Med.*, Vol. 1, pp. 51–58.

SIEBERT, W.M. (1986), *Circuits, signals and systems*, The MIT Press, Massachusetts.

SILVER, S. (1949), *Microwave antenna theory and design*, McGraw Hill, New York, 1st ed.

SINHA, N.K. and KUSZTA, B. (1983), *Modelling and identification of dynamic systems*, Van Nostrand Reinhold Co, New York.

SKOLNIK, M.I. (1980), *Introduction to RADAR systems*, McGraw Hill, New York.

SLACK, C.W. (1953), 'Learning in simple one-dimensional tracking', *Amer. J. Psychol.*, Vol. 66, pp. 33–43.

SMITH, F.G. (1966), *Radio astronomy*, Penguin Books, Harmondworth, England, 3rd ed.

STEINHARDT, A.O. (1988), 'Householder transforms in signal processing', *IEEE Acoustics Speech Signal Process. Magazine*, Vol. 5, No. 3, pp. 4–12.

STOCKHAM, T.G., CANNON, T.M. and INGEBRETSEN, R.B. (1975), 'Blind deconvolution through digital signal processing', *Proc. IEEE*, Vol. 63, pp. 678–692.

STRANG, G. (1980), *Linear algebra and its applications*, Academic Press, London, 2nd ed.

TEKALP, A.M., KAUFMAN, H. and WOODS, J.W. (1986), 'Identification of image and blur parameters for the restoration of noncausal blurs', *IEEE Trans Acoustics Speech Signal Process.*, Vol. 34, pp. 963–972.

TITCHMARSH, E.C. (1939), *The theory of functions*, Oxford University Press, London, second ed.

TODD-POKROPEK, A. (1980), 'Image processing in nuclear medicine', *IEEE Trans. Nucl. Med.*, Vol. 27, No. 3, pp. 1080–1093.

TODOSIEV, E.P., ROSE, R.E. and SUMMERS, L.G. (1967), 'Human performance in single and two-axis tracking systems', *IEEE Trans. Hum. Factors in Electron.*, Vol. 8, pp. 125–129.

TRIER ANDERSEN, O. (1986), 'A system for quantitative assessment of dyscoordination and tremor', *Acta Neurol. Scand.*, Vol. 73, pp. 291–294.

TRUSSELL, H.J. (1983), 'Convergence criteria for iterative restoration methods', *IEEE Trans. Acoustics Speech Signal Process.*, Vol. 31, pp. 129–136.

URICK, R.J. (1967), *Principles of underwater sound*, McGraw Hill, New York.

VIVIANI, P. (1988), 'Two-dimensional visuo-manual tracking: beyond the system analysis approach', In *Proc. 1988 IEEE Int. Conf. Syst. Man Cyber.*, International Academic, Beijing.

VIVIANI, P. and MOUNOUD, P. (1990), 'Perceptuomotor compatibility in pursuit tracking of two-dimensional movements', *J. Motor Behav.*, Vol. 22, pp. 407–443.

VIVIANI, P., CAMPADELLI, P. and MOUNOUD, P. (1987), 'Visuo-manual pursuit tracking of human two-dimensional movements.', *J. Exp. Psychol.*, Vol. 13, pp. 62–78.

WALKER, R.J. (1950), *Algebraic curves*, Princeton University Press, New Jersey.

WARREN, C.E., FITTS, P.M. and CLARK, J.R. (1952), 'An electronic apparatus for the study of the human operator in a 1-dimensional closed-loop continuous pursuit task', *AIEE Trans.*, Vol. 71, pp. 19–23.

WATSON, R.W. and BONES, P.J. (1991), 'An alternative to Wiener filtering', *Proceedings of the 6th N.Z. Image Processing Workshop*, pp. 139–144.

WATSON, R.W. and BONES, P.J. (1992), 'An alternative to Wiener filtering II', In BONES, P. (Ed.), *Proceedings of the 7th N.Z. Image Processing Workshop*, pp. 71–76.

WATSON, R., PARKER, C. and BONES, P. (1992), 'Demonstration of 2-D consistent deconvolution using zeros', *Optics Comm.*, Vol. 93, pp. 359–365.

WELFORD, A.T. (1968), *Fundamentals of skill*, Methuen and Company, London.

WESTLAKE, J.R. (1968), *A handbook of numerical matrix inversion and solution of linear equations*, Wiley, New York.

WHITING, H.T.A. (1975), *Concepts in skill learning*, Lepus Books, London.

WILLIGES, R.C. and WILLIGES, B.H. (1978), 'Critical variables in adaptive motor skills training', *Human Factors*, Vol. 20, pp. 201–214.

WITTEN, I.M. (1982), *Principles of computer speech*, Academic Press, London.

WOODS, J.W. and INGLE, V.K. (1981), 'Kalman filtering in two dimensions: further results', *IEEE Trans Acoustics Speech Signal Process.*, Vol. 29, pp. 188–197.

WOODS, J.W. and RADEWAN, C.H. (1977), 'Kalman filtering in two dimensions', *IEEE Trans Inform. Theory*, Vol. 23, pp. 473–482.

WU, Q.X., BONES, P.J. and BATES, R.H.T. (1989), 'Translational motion compensation for coronary angiogram sequences', *IEEE Trans. Med. Imaging*, Vol. 8, pp. 276–282.

YOULA, D.C. (1987), 'Mathematical theory of image restoration by the method of convex projections', In STARK, H. (Ed.), *Image recovery theory and applications*, Academic Press, London.

YOULA, D.C. and WEBB, H. (1982), 'Image reconstruction by the method of convex projections: Part 1 - theory', *IEEE Trans. Med. Imaging*, Vol. 1, pp. 81–94.

ZAKHOR, A. and IZRAELEVITZ, D. (1986), 'A note on the sampling of zero crossings of two-dimensional signals', *Proc. IEEE*, Vol. 74, pp. 1285–1287.

ZAKHOR, A. and OPPENHEIM, A.V. (1990), 'Reconstruction of two-dimensional signals from level crossings', *Proc. IEEE*, Vol. 78, pp. 31–55.

ZHUANG, X., ØSTEVOLD, E. and HARALICK, R.M. (1987), 'The principle of maximum entropy in image recovery', In STARK, H. (Ed.), *Image recovery theory and applications*, Academic Press, London.

ZIEGLER, P.N. (1968), 'Single and dual axis tracking as a function system dynamics', *Hum. Factors*, Vol. 10, pp. 273–275.